

GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2002, 20:26:02 ; Search time 1873.74 Seconds
(without alignments)
13513.673 Million cell updates/sec

Title: US-09-805-919-3
Perfect score: 1210
Sequence: 1 tatattcaattgaaacatg.....ttcgcaaaaaaaaaaaaaa 1210

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 1797656 seqs, 10463268293 residues
Total number of hits satisfying chosen parameters: 3595312

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenBank
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_inv:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No. Score Query Match length DB ID Description

C	1	58.8	4.9	611	6	AX187051	AX187051 Sequence
C	2	55.6	4.6	33014	3	CER06B9	Z83237 Caenorhabdi
C	3	55.4	4.6	7736	3	AF310889	AF310889 Dictyoste
C	4	55.4	4.6	63807	2	AC087299	AC087299 Homo sapi
C	5	55.4	4.6	11328	9	AL390997	AL390997 Human DNA
C	6	55.4	4.6	147891	2	AL365191	AL365191 Homo sapi
C	7	55.4	4.6	183892	2	AC019074	AC019074 Homo sapi
C	8	55.4	4.6	216680	2	AC021211	AC021211 Homo sapi
C	9	55.2	4.6	19209	1	AF274444	AF274444 Carsonell
	10	54	4.5	11662	6	AX346802	AX346802 Sequence
	11	53.8	4.4	234112	3	PFMAL4P2	AL035475 Plasmodiu
	12	53.4	4.4	8952	6	AX251200	AX251200 Sequence
	13	53.4	4.4	17869	6	AX345007	AX345007 Sequence
	14	53.4	4.4	17869	6	AX348307	AX348307 Sequence
	15	53	4.4	198	6	AX182091	AX182091 Sequence
	16	53	4.4	169546	2	AC004157	AC004157 Plasmodiu
	17	52.8	4.4	110000	2	PFMAL4P1_2	Continuation (3 of
	18	52.6	4.3	318221	2	PFMAL13P3	AL049184 Plasmodiu
	19	52.4	4.3	425	6	AX284646	AX284646 Sequence
	20	52.4	4.3	1462	10	BC006018	BC006018 Mus muscu
	21	52.4	4.3	8805	6	AX345415	AX345415 Sequence
	22	52.4	4.3	14920	6	AX344738	AX344738 Sequence
	23	52.2	4.3	46353	2	AC100712	AC100712 Mus muscu
	24	52.2	4.3	12592	6	AX251796	AX251796 Sequence
	25	52.2	4.3	14950	6	AX346159	AX346159 Sequence
	26	52.2	4.3	170941	2	AC068027	AC068027 Homo sapi
	27	52.2	4.3	172805	9	AL360270	AL360270 Human DNA
	28	52.2	4.3	181936	2	AC013387	AC013387 Homo sapi
	29	52.2	4.3	184706	2	AC013327	AC013327 Homo sapi
	30	52.2	4.3	186462	2	AL391727	AL391727 Homo sapi
	31	52	4.3	6325	2	DDI012088	Continuation (4 of
	32	51.8	4.3	77835	2	PFMAL13P2_3	AX345403 Sequence
	33	51.6	4.3	8342	6	AX345403	AC005506 Plasmodiu
	34	51.6	4.3	205429	2	AC005506	AE001391 Plasmodiu
	35	51.2	4.2	14553	3	AE001391	AX251551 Sequence
	36	51.2	4.2	56153	6	AC023330	AC023330 Homo sapi
	37	51.2	4.2	81358	2	AC023330	U87514 Dictyostell
	38	51	4.2	975	3	DDU87514	AE001372 Plasmodiu
	39	51	4.2	12029	3	AE001372	AC106404 Rattus no
	40	51	4.2	157373	2	AC106404	AX346543 Sequence
	41	50.8	4.2	5449	6	AX346543	AX278022 Sequence
	42	50.8	4.2	7657	6	AX278022	AX346924 Sequence
	43	50.8	4.2	7657	6	AX346924	AX344729 Sequence
	44	50.8	4.2	11209	6	AX344729	AX281321 Sequence
	45	50.8	4.2	13123	6	AX281321	

ALIGNMENTS

RESULT 1
AX187051/c
LOCUS AX187051 611 bp DNA
DEFINITION Sequence 2746 from Patent WO0142467.
ACCESSION AX187051
VERSION AX187051.1 GI:15138495
KEYWORDS
SOURCE human.
ORGANISM Homo sapiens
REFERENCE
AUTHORS Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
TITLE 1 (bases 1 to 611)
JOURNAL Genes, compositions, kits, and methods for identification, assessment, prevention, and therapy of cervical cancer
Patent: WO 0142467-A 2746 14-JUN-2001;
FEATURES
source
1. 611
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"

BASE COUNT 379 a 50 c 2 g 178 t 2 others
ORIGIN

gene	OFLPAPVYPQDPNLPQILPTPIGGVAVAPPQYGFLLPMADGSEKAGADASYPAASMP YPIPNVPYPSKDDGTIVVIPSAPPLSYEESMEGQDGTETLNTDENQDPAKYPVFNNLP VYNPTAPPE"					
	join(27437. .27621,27944. .28049,28096. .28270,28695. .29366, 29417. .29562)	/gene="R06B9.1"				
CDS	join(27437. .27621,27944. .28049,28096. .28270,28695. .29366, 29417. .29562)	/gene="R06B9.1" /note="Similarity to Human dihydroxyvitamin D3-induced protein" /codon_start=1 /protein_id="CAB05786.1" /db_xref="GI:3878864" /db_xref="SPTREMBL:O17981" /translation="MPETELHVIFDQPNVEFFPGQISGRVVLSTTEKKARAVNIK ILGLAHTSWTDYDSVRVDNDGNVRYDSESVHSSNVHYLDQALLIMACKDGSNELSA GDYWPFSYTLPLNPPSFEFGKGYLRYSVAEVDPRWLDKAKRCITGEARLSPLI DLNAIPLALTPIDDESEENLGCCFFRKGYLELRVNIPTGFVPGETVPMNITHLNSS VPVTEVKAKIIQCKFTAYRNGTIFRFDGSDTLMSSSQCKYDTKPVITQTQPMYV TPGNEHKEFVLEFRLPSVPTPTICRFSPVITVEYVQVRVETTSICGSAKCEMPLIGT VPIRMYLPPIPNMYPIGLPPPYANLTDVNVCPSGSGSITAVIPSAPPPSYQESMYGV GGTELRAEENEKPFAPKYPVENNLPYNNPTAPPE" join(30328. .30503,31116. .31221,31318. .31483,32023. .32123, 32186. .32413) /gene="R06B9.2" join(30328. .30503,31116. .31221,31318. .31483,32023. .32123, 32186. .32413) /gene="R06B9.2" /note="Similarity to Human dihydroxyvitamin D3-induced protein" /codon_start=1 /protein_id="CAB05789.1" /db_xref="GI:3878867" /db_xref="SPTREMBL:O17984" /translation="MPELFLVLDKPNNAVYAAGOKISGRVVFSTASQONPRMIDVOLHG RSHFTFTROSEETKINSKGESETKTHVHYTATAKHLDTAVPLMRKTDKARLLPGKY EMQFWQLPCSVLPPSEFGNNGNIRYWRVAEVSRSKMNIVDESSFEIAPFLDLNTP IARTPLDGFVKNLGGCCFRNNGNSVSDVQTVHDIVHSVRNTVKDGLAKILLQRSQEN GNIIQGGILNGFRERGIHDFGLEGLLNKGSDHGTDYIFG"				
gene	join(30328. .30503,31116. .31221,31318. .31483,32023. .32123, 32186. .32413)				/gene="R06B9.2"	
CDS	join(30328. .30503,31116. .31221,31318. .31483,32023. .32123, 32186. .32413)				/gene="R06B9.2"	
	/note="Similarity to Human dihydroxyvitamin D3-induced protein"					
	/codon_start=1					
	/protein_id="CAB05789.1"					
	/db_xref="GI:3878867"					
	/db_xref="SPTREMBL:O17984"					
	/translation="MPELFLVLDKPNNAVYAAGOKISGRVVFSTASQONPRMIDVOLHG RSHFTFTROSEETKINSKGESETKTHVHYTATAKHLDTAVPLMRKTDKARLLPGKY EMQFWQLPCSVLPPSEFGNNGNIRYWRVAEVSRSKMNIVDESSFEIAPFLDLNTP IARTPLDGFVKNLGGCCFRNNGNSVSDVQTVHDIVHSVRNTVKDGLAKILLQRSQEN GNIIQGGILNGFRERGIHDFGLEGLLNKGSDHGTDYIFG"					
BASE COUNT	10605 a 6288 c 5920 g 10201 t					
ORIGIN						
Query Match	4.6%; Score 55.6; DB 3; Length 33014;					
Best Local Similarity	52.1%; Pred. No. 0.047;					
Matches 124; Conservative	0; Mismatches 114; Indels 0; Gaps 0;					
QY 939	tggttaagattttctgtagtcaagctattttaacagaagttctgtattttttccagaaaa 998					
Db 10619	TTTTAAACATTTTAAACAAGTTTTTTTGAATTTTAAATTTTGTGAAAT 10560					
QY 999	tctaggacaagtcacacctgtgctggcgattaactactaggattttcttccagtttag 1058					
Db 10559	TCTTAGAATCTTTCGACCTTTTCTTGAATTTTCTTAACCTTTTTCGAAATTAATTG 10500					
QY 1059	tccgtgattttatttgatatcttaccatttgattgtgtatgattttttccctaaaaat 1118					
Db 10499	AATTTTGTGTTTGAATTAATAATTTTTCGAATATTGGGGATTTTTGAACGAAAAAT 10440					
QY 1119	tttataatttcctaattctgtlaagtaattgaatgatatgttacttctgtccaat 1176					
Db 10439	TCGAAATTTTGTAGAGATTTTGTAGGGAATTTTGAAGTGAATTCGAATTTTTCGGTGAT 10382					
RESULT 3						
AF310889	AF310889	7736 bp	DNA linear INV 06-MAR-2001			
LOCUS						
DEFINITION	Dictyostelium discoideum RacD (racd) gene, complete cds; and unknown genes.					
ACCESSION	AF310889					
VERSION	AF310889.1 GI:12007300					
KEYWORDS						

SOURCE	Dictyostelium discoideum.
ORGANISM	Dictyostelium discoideum
REFERENCE	Eukaryota; Mycetozoa; Dictyostelida; Dictyostelium.
AUTHORS	1 (bases 1 to 7736)
TITLE	Rivero,F., Dislich,H., Glockner,G. and Noegel,A.A.
JOURNAL	The Dictyostelium discoideum family of Rho-related proteins
MEDLINE	Nucleic Acids Res. 29 (5), 1068-1079 (2001)
REFERENCE	21127961
AUTHORS	2 (bases 1 to 7736)
TITLE	Rivero,F., Dislich,H. and Noegel,A.A.
JOURNAL	Direct Submission
FEATURES	Submitted (02-OCT-2000) Institut fuer Biochemie I, Medizinische Fakulttaet, Universitaet zu Koeln, Joseph-Stelzmann-St. 52, Cologne 50931, Germany
source	location/Qualifiers
1. .7736	
/organism="Dictyostelium discoideum"	
/strain="AX4"	
/db_xref="taxon:44689"	
join(<856. .1007,1160. .>4670)	
/product="unknown"	
join(856. .1007,1160. .4670)	
/note="5' region identified manually, therefore unsure; similarity to mouse SH3-domain containing adaptor protein (central part) and rat L-glutamine aminohydrolase (C-terminal part)"	
/codon_start=1	
/product="unknown"	
/protein_id="AAG45120.1"	
/db_xref="GI:12007301"	
/translation="MFKWQSNLNVYTFENIPANVTKODIHAETSSSIGGVKGFGL NCEVCIFIKGSKWAIKDDTVQVQMDKQQTGVKWSHLSEITESHNAKAVLFEVATSE EELSLIPHQVIDIISKDDSGWEGSKQOMGVFPSPNQEFPTDYQPPEDDASAEAT KEEIIKSPIQSVRLIGINISLGSMDLKNKLAIRKTYATTSADIIEAHNNNANSEES NSSSSSSSSNNSSGGFVAVKPSGAVNTQQRQPKVILNNROSTILSGVPAIQISIP SSNTNTTNTTTPAVOPLQLKPKQNTTEINTSGNSIPATTSKESOSTKESSSGN KITSLFKKPLKKITAISALTGSHSNSSSGGDKNNSGSKTAKAVYDFESKEPNELN LKKGDIIVVLAKDSGWMOGINOSTQATGWFNTFVEELKEEKLPPSSSENKOPLK IVKSPPKSSSSSQPQTEAKLQSLSDTPKLETARKAGAAGRKPPSRVRSYLLSD EKAEREKIKOSFYSHVSSOKELSMYDNNEDGHEHSDDOQSTPSPISRSGSFTPT QQQQQSVDKPIERPSKPPPPKRTAAMNSSLNINNEPPPSNKKPPPPKSLKPLSTSD GYEPPSPSSSTPTITTSGEDTLOKSPLOKRVAPPOQONEDLPSSLVGKPPPLK PTPKPRAIQTNSESTAPAPTPSHEYNFKANLPAKPPSSPIGNSQEIPTDISSPPNAT TPQTPPPLKRYAKPPVOQOQPSNENIENTOFKLKPSAPSSPSPIDISTTTK PPARPPLKKSAAAAAPSPSPNEEATTTSPPTLAKPKVALRNGSSAPTSLSQPOQP LPSTHKRNLSNPTNTNNEKPPPTKKPOLIKPLPOFNDCVEVSNICNEIQERSDNS QETTTFKHNPSSLIDSEFSVAVCNDGQFWGNDQLDHCVPMQVITYPFLYSILCE EIGLDEVSKFISDOARPDSDYDNOKSLNDOKKSHNPFTLAGOLITANLSLVNIMHYLE STQFLNITQELVNCSSVSCDMVSIYSOKSETSDIELVGHLLKSSGLIKQPDVLDFFY QINSIOLNSKOVSVALATLANDGVPESODLKLAPKNIISKTTDLIRICNSTPTTSSI FDNNIDNINNSKIIPILSDSGVLMIIIPGLMGVSIISPCNNNNNNNNNDNFNHYPN NNHIEFCRKLCSILDQ"	
join(5116. .5180,5293. .5333,5523. .5590,5805. .5914, 6023. .6634)	
/gene="racd"	
/note="corresponds to EST clones SSB614, SSG550, SSJ204 and SSK792 from the Japanese cDNA sequencing project; cDNA (partial) already described in L11594."	
/product="RacD"	
5116. .6634	
/gene="racd"	
join(5163. .5180,5293. .5333,5523. .5590,5805. .5914, 6023. .6550)	
/gene="racd"	
/note="Rho-GTPase; lacks prenylation motif characteristic of Rho proteins"	
/codon_start=1	
/product="RacD"	
/protein_id="AAG45121.1"	
/db_xref="GI:12007302"	
/translation="MATGIKKTYKVVVVGDAVGKTSLLILYTTKAFPKDYVPTVEDN FNCLEMYDNKPVNLVLMDTAGQEDYDNLRLPSYPQTDVFIICYSVVKRDSLDNIKYKW LPEINQFNQGTPIIILVGTKTDLRBDKKTLLSQLESKQEPVSRDEGVALAKEIGAVQGF	

18404 19087: contig of 684 bp in length
* 19088 19187: gap of 100 bp
* 19188 19882: contig of 695 bp in length
* 19883 19982: gap of 100 bp
* 19983 20676: contig of 694 bp in length
* 20677 20776: gap of 100 bp
* 20777 21473: contig of 697 bp in length
* 21474 21573: gap of 100 bp
* 21574 22282: contig of 709 bp in length
* 22283 22382: gap of 100 bp
* 22383 23092: contig of 710 bp in length
* 23093 23192: gap of 100 bp
* 23193 23872: contig of 680 bp in length
* 23873 23972: gap of 100 bp
* 23973 24731: contig of 759 bp in length
* 24732 24831: gap of 100 bp
* 24832 25530: contig of 699 bp in length
* 25531 25630: gap of 100 bp
* 25631 26333: contig of 703 bp in length
* 26334 26433: gap of 100 bp
* 26434 27122: contig of 689 bp in length
* 27123 27222: gap of 100 bp
* 27223 27894: contig of 672 bp in length
* 27895 27994: gap of 100 bp
* 27995 28670: contig of 676 bp in length
* 28671 28770: gap of 100 bp
* 28771 29478: contig of 708 bp in length
* 29479 29578: gap of 100 bp
* 29579 30282: contig of 704 bp in length
* 30283 30382: gap of 100 bp
* 30383 31081: contig of 699 bp in length
* 31082 31181: gap of 100 bp
* 31182 31853: contig of 672 bp in length
* 31854 31953: gap of 100 bp
* 31954 32660: contig of 707 bp in length
* 32661 32760: gap of 100 bp
* 32761 33474: contig of 714 bp in length
* 33475 33574: gap of 100 bp
* 33575 34291: contig of 717 bp in length
* 34292 34391: gap of 100 bp
* 34392 35072: contig of 681 bp in length
* 35073 35172: gap of 100 bp
* 35173 35857: contig of 685 bp in length
* 35858 35957: gap of 100 bp
* 35958 36627: contig of 670 bp in length
* 36628 36727: gap of 100 bp
* 36728 37438: contig of 711 bp in length
* 37439 37538: gap of 100 bp
* 37539 38248: contig of 710 bp in length
* 38249 38348: gap of 100 bp
* 38349 39043: contig of 695 bp in length
* 39044 39143: gap of 100 bp
* 39144 39842: contig of 699 bp in length
* 39843 39942: gap of 100 bp
* 39943 40657: contig of 715 bp in length
* 40658 40757: gap of 100 bp
* 40758 41433: contig of 676 bp in length
* 41434 41533: gap of 100 bp
* 41534 42176: contig of 643 bp in length
* 42177 42276: gap of 100 bp
* 42277 42985: contig of 709 bp in length
* 42986 43085: gap of 100 bp
* 43086 43793: contig of 708 bp in length
* 43794 43893: gap of 100 bp
* 43894 44609: contig of 716 bp in length
* 44610 44709: gap of 100 bp
* 44710 45405: contig of 696 bp in length
* 45406 45505: gap of 100 bp
* 45506 46211: contig of 706 bp in length
* 46212 46311: gap of 100 bp
* 46312 46988: contig of 677 bp in length
* 46989 47089: gap of 100 bp
* 47089 47802: contig of 714 bp in length

* 47803 47902: gap of 100 bp
* 47903 48614: contig of 712 bp in length
* 48615 48714: gap of 100 bp
* 48715 49436: contig of 722 bp in length
* 49437 49536: gap of 100 bp
* 49537 50219: contig of 683 bp in length
* 50220 50319: gap of 100 bp
* 50320 50991: contig of 672 bp in length
* 50992 51091: gap of 100 bp
* 51092 51777: contig of 686 bp in length
* 51778 51877: gap of 100 bp
* 51878 52599: contig of 722 bp in length
* 52600 52699: gap of 100 bp
* 52700 53420: contig of 721 bp in length
* 53421 53520: gap of 100 bp
* 53521 54215: contig of 695 bp in length
* 54216 54315: gap of 100 bp
* 54316 55020: contig of 705 bp in length
* 55021 55120: gap of 100 bp

Query Match 4.6%; Score 55.4; DB 2; Length 63807;
Best Local Similarity 49.8%; Pred. No. 0.051;
Matches 140; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 928 atgaacagaactgttaagatttttctagtcgaagctattttaagaagttgtgtattt 987
DB 34594 AAGAAATAGGACCTAGAAATTTGCTAGCACAGCAGGCTAACTATAGCAAAATATATTATA 34653
QY 988 ttccagaaatctagacaagtcgaacctgtctgcgcatctaactagatttctt 1047
DB 34654 ATTCTAGGTATATATATTTTATGTGTGGCTGTGTTAAATAGGATTAACGTTAATTCTTT 34713
QY 1048 ttccagtttagtcctgtatttatttgatattcttaccatttgattgtgtatgtt 1107
DB 34714 TTTCAGATTGTTTCATAGATGCTATATAGAAATGCTACTGATTGTGAGTGTGATTGTGT 34773
QY 1108 ttccctaaatttataatttcctaattcttgaagtgaattgaattgataattgttactt 1167
DB 34774 ATCTGAAACTTTAATGTAATTTATCATCAGTCTTAATAGTTTGTGAGGATTTCAAGGTTT 34833
QY 1168 tctgtcaataatagaacaagacattcgcaaaaaaaaaa 1208
DB 34834 TTTTCAATATATAGATCATCATCTGCAACGATTAATTAA 34874

RESULT 5
AL390997/c
LOCUS
DEFINITION Human DNA sequence from clone RP11-159G6 on chromosome Xq21.33-23, complete sequence.
ACCESSION AL390997
VERSION AL390997.10 GI:10432544
KEYWORDS HTG.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 112328)
AUTHORS Howden, P.
TITLE Direct Submission
JOURNAL Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
COMMENT Submitted (06-DEC-2001) Wellcome Trust Sanger Institute, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquerry@sanger.ac.uk
On Sep 29, 2000 this sequence version replaced gi:10303411.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em.; EMBL; Sw.; SWISSPROT; Tr.; TREMBL; Wp.; WORMPEP; Information

on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human Chromosome X, constructed by the Sanger Centre Chromosome X Mapping Group. Further information can be found at
<http://www.sanger.ac.uk/HGP/ChrX>

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. RP11-159G6 is from the library RPCI-11.1 constructed by the group of Pieter de Jong. For further details see
<http://www.chori.org/bacpac/home.htm>

VECTOR: pBACE3.6

IMPORTANT: This sequence is not the entire insert of clone RP11-159G6. It may be shorter because we sequence overlapping sections only once, except for a short overlap.

The true right end of clone RP11-159G6 is at 112328 in this sequence. The true right end of clone RP13-130D24 is at 100 in this sequence. The true right end of clone RP11-115M20 is at 27343 in this sequence.

FEATURES

```
source
1. .112328
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/map="q21.33-23"
/clone_lib="RP11-159G6"
/clone_lib="RPCI-11.1"
24. .4692
repeat_region
/note="LIM1 repeat: matches -1390. .3476 of consensus"
4704. .5475
repeat_region
/note="LIMC3 repeat: matches 6280. .7003 of consensus"
5489. .5609
repeat_region
/note="MER5A repeat: matches 1. .121 of consensus"
8421. .8505
repeat_region
/note="MIR repeat: matches 164. .244 of consensus"
10380. .11089
repeat_region
/note="MER21B repeat: matches 37. .713 of consensus"
MER21B repeat: matches 37. .713 of consensus"
11416. .11644
repeat_region
/note="LIMEc repeat: matches 1643. .1875 of consensus"
11474. .11756
repeat_region
/note="HAL1 repeat: matches 625. .912 of consensus"
12125. .12487
repeat_region
/note="LIM4 repeat: matches 2148. .2523 of consensus"
12581. .12850
repeat_region
/note="LIM4 repeat: matches 2857. .3140 of consensus"
12911. .12958
repeat_region
/note="MST-INTERNAL repeat: matches 1586. .1635 of consensus"
MST-INTERNAL repeat: matches 1586. .1635 of consensus"
12966. .13057
repeat_region
/note="MLT2D repeat: matches 3. .102 of consensus"
13112. .13459
repeat_region
/note="MST-INTERNAL repeat: matches 1233. .1587 of consensus"
MST-INTERNAL repeat: matches 1233. .1587 of consensus"
13463. .13575
repeat_region
/note="MLT2D repeat: matches 3. .122 of consensus"
13650. .14015
repeat_region
/note="AluDo repeat: matches 3. .239 of consensus"
14146. .14224
repeat_region
/note="LIMC repeat: matches 1788. .1863 of consensus"
14225. .14537
repeat_region
/note="AluDo repeat: matches 1. .312 of consensus"
14538. .15365
repeat_region
/note="LIMEc repeat: matches 1863. .2764 of consensus"
15384. .17055
repeat_region
/note="LIM1 repeat: matches -1390. .269 of consensus"
```

```
repeat_region
/note="L1 repeat: matches 2851. .4751 of consensus"
18982. .19698
repeat_region
/note="LIM2 repeat: matches 4957. .5674 of consensus"
19716. .19775
repeat_region
/note="L1 repeat: matches 3926. .3986 of consensus"
19786. .20381
repeat_region
/note="LIM2 repeat: matches 5681. .6308 of consensus"
20484. .20987
repeat_region
/note="LIMB7 repeat: matches 2802. .3311 of consensus"
20988. .21046
repeat_region
/note="MADE1 repeat: matches 1. .80 of consensus"
21047. .23080
repeat_region
/note="LIMB7 repeat: matches 3311. .5305 of consensus"
21489. .21554
repeat_region
/note="33 copies 2 mer ta 71% conserved"
23108. .26375
repeat_region
/note="LIP47 repeat: matches 2872. .6145 of consensus"
23110. .26375
repeat_region
/note="LIP48 repeat: matches 2872. .6161 of consensus"
26376. .27185
repeat_region
/note="LIMB7 repeat: matches 5302. .6131 of consensus"
27279. .27880
repeat_region
/note="L1 repeat: matches 3729. .4348 of consensus"
28069. .28140
repeat_region
/note="LIMB3 repeat: matches 6109. .6180 of consensus"
28173. .29491
repeat_region
/note="LIP413 repeat: matches 4812. .6156 of consensus"
32110. .32514
repeat_region
/note="LIMC4 repeat: matches 6513. .6938 of consensus"
32573. .32807
repeat_region
/note="LIMC4 repeat: matches 7026. .7288 of consensus"
33293. .34392
repeat_region
/note="LIP47 repeat: matches 5025. .6141 of consensus"
LIP47 repeat: matches 5025. .6141 of consensus"
34555. .36545
repeat_region
/note="LIPB2 repeat: matches 4140. .6155 of consensus"
LIPB2 repeat: matches 4140. .6155 of consensus"
36579. .36893
repeat_region
/note="MLT1F repeat: matches 180. .510 of consensus"
37365. .37532
repeat_region
/note="MIR repeat: matches 48. .225 of consensus"
38511. .38637
repeat_region
/note="LIME repeat: matches 5677. .5802 of consensus"
38710. .38847
repeat_region
/note="MLT1J repeat: matches 109. .237 of consensus"
39170. .39353
repeat_region
/note="MIR repeat: matches 3. .210 of consensus"
40495. .40572
repeat_region
/note="L2 repeat: matches 2620. .2710 of consensus"
42209. .42346
repeat_region
/note="MER5A repeat: matches 1. .139 of consensus"
42362. .42804
repeat_region
/note="LIP416 repeat: matches 5709. .6157 of consensus"
42805. .42866
repeat_region
/note="MER5A repeat: matches 123. .186 of consensus"
43090. .43703
repeat_region
/note="L2 repeat: matches 2156. .2749 of consensus"
44064. .44723
repeat_region
/note="LIP49 repeat: matches 5498. .6160 of consensus"
44066. .44709
repeat_region
/note="LIP410 repeat: matches 5510. .6158 of consensus"
44724. .44771
repeat_region
/note="AluSp/q repeat: matches 256. .303 of consensus"
44772. .44908
repeat_region
/note="LIP410 repeat: matches 5391. .5527 of consensus"
44775. .44906
repeat_region
/note="LIP49 repeat: matches 5393. .5524 of consensus"
44908. .46912
repeat_region
/note="LIP repeat: matches 3080. .5042 of consensus"
LIP repeat: matches 3080. .5042 of consensus"
46913. .47229
repeat_region
/note="AluSq repeat: matches 1. .310 of consensus"
```

```
repeat_region 47230..47530
/note="LIP repeat: matches 5042..5347 of consensus
LIP repeat: matches 5042..5347 of consensus"
repeat_region 47527..47654
/note="LIP repeat: matches 3312..3439 of consensus
LIP repeat: matches 3312..3439 of consensus"
repeat_region 47902..50241
/note="LIMB1 repeat: matches 3808..6155 of consensus
LIMB1 repeat: matches 3808..6155 of consensus"
repeat_region 50245..50593
/note="L1 repeat: matches 3024..3367 of consensus"
repeat_region 50594..50949
/note="MLT1A1 repeat: matches 1..365 of consensus"
repeat_region 50950..51819
/note="L1 repeat: matches 2131..3024 of consensus"
repeat_region 51822..52183
/note="L1PB1 repeat: matches 5781..6155 of consensus"
repeat_region 52287..53414
/note="L1PB1 repeat: matches 4656..5791 of consensus"
repeat_region 53426..53559
/note="L1PA15 repeat: matches 3620..3753 of consensus
L1PA15 repeat: matches 3620..3753 of consensus"
repeat_region 53527..55739
/note="L1PA15 repeat: matches 3904..6152 of consensus
L1PA15 repeat: matches 3904..6152 of consensus"
repeat_region 55746..55803
/note="29 copies 2 mer aa 70% conserved"
repeat_region 56148..56193
```

Query Match 4.6%; Score 55.4; DB 9; Length 112328;
Best Local Similarity 49.8%; Pred. No. 0.05;
Matches 140; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

```
QY 928 atgacagaactgttaagattttgtgctgaagctatttaacagaagttgtgattt 987
Db 18704 AAGAATAGGACCTAGATTGCTAGACACAGGTAACCTAGTCAAAAATAATTATA 18645

QY 988 tttcagaaatctagacaagctgtgctgctgcgattactaattagattttct 1047
Db 18644 ATTCTAGGTATATAATTTATGTGTGCTGTTGTAATAAGATTAACGTTAATTCCTT 18585

QY 1048 ttccagttagtcctgtattttatttgatattcttacctatttgattgtgtatgatttt 1107
Db 18584 TTTCAGATTGTTCAATAGATGATATAGAAGTCTACTGATTTGTGAGTGTGATTGT 18525

QY 1108 ttccctaaatttataatttcctcaattctgtgaatgaatgatatattgtaatt 1167
Db 18524 ATCCTGAACCTTATGATTTATCAGTCTCAATAGTTTGTGTGAGTATTCAGTCTT 18465

QY 1168 tctgcaataatagacaagacattgcacaaaaaa 1208
Db 18464 TTTTCAATATATAAGATCACATCATCTGCAACGATAATTAA 18424
```

RESULT 6
AL365191/c 147891 bp DNA linear HTG 10-JUL-2001
LOCUS Homo sapiens chromosome X clone RP11-115M20, *** SEQUENCING IN
DEFINITION PROGRESS ***, 18 unordered pieces.
ACCESSION AL365191
VERSION AL365191.4 GI:10129532
KEYWORDS HTG; HTGS_PHASE1; HTGS_CANCELLED.
SOURCE human.
ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 147891)

AUTHORS McIay,K.
TITLE Direct Submission
JOURNAL Submitted (09-JUL-2001) Sanger Centre, Hinxton, Cambridgeshire,
CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk Clone
requests: clonerequests@sanger.ac.uk

COMMENT On Sep 14, 2000 this sequence version replaced gi:9801124.

----- Genome Center
Center: Sanger Centre
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: BA115M20
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Sequencing vector: plasmid; L08752; 100% of reads
Chemistry: Dye-terminator Big Dye; 52% of reads
Chemistry: Dye-terminator ET-amersham; 47% of reads
Consensus quality: 14347 bases at least Q40
Consensus quality: 14976 bases at least Q30
Insert size: 14619; sum-of-contigs
Insert size: 180235; 9.5% error; agarose-fp
Quality coverage: 4.35x in Q20 bases; sum-of-contigs quality
coverage: 3.71x in Q20 bases; agarose-fp

* NOTE: This is a 'working draft' sequence. It currently
* consists of 18 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```
1 4096: contig of 4096 bp in length
* 4097 4196: gap of 100 bp
* 4197 6285: contig of 2089 bp in length
* 6286 6385: gap of 100 bp
* 6386 14509: contig of 8124 bp in length
* 14510 14609: gap of 100 bp
* 14610 17060: contig of 2451 bp in length
* 17061 17160: gap of 100 bp
* 17161 21976: contig of 4816 bp in length
* 21977 22076: gap of 100 bp
* 22077 30708: contig of 8632 bp in length
* 30709 30808: gap of 100 bp
* 30809 46192: contig of 15384 bp in length
* 46193 46292: gap of 100 bp
* 46293 89106: contig of 42814 bp in length
* 89107 89206: gap of 100 bp
* 89207 97724: contig of 8518 bp in length
* 97725 97824: gap of 100 bp
* 97825 103968: contig of 6144 bp in length
* 103969 104068: gap of 100 bp
* 104069 106312: contig of 2244 bp in length
* 106313 106412: gap of 100 bp
* 106413 110483: contig of 4071 bp in length
* 110484 110583: gap of 100 bp
* 110584 117466: contig of 6883 bp in length
* 117467 117566: gap of 100 bp
* 117567 127508: contig of 9942 bp in length
* 127509 127608: gap of 100 bp
* 127609 130184: contig of 2576 bp in length
* 130185 130284: gap of 100 bp
* 130285 132386: contig of 2102 bp in length
* 132387 132486: gap of 100 bp
* 132487 141197: contig of 8711 bp in length
* 141198 141297: gap of 100 bp
* 141298 147891: contig of 6594 bp in length.
```

FEATURES
source 1..147891

/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="X"
/clone_lib="RP11-115M20"
/clone_id="RP11-11.1"
1..4096
/note="assembly_fragment:00554
fragment_chain:1

misc_feature

```

clone_end:T7
vector_side:left"
4197. .6285
/note="assembly_fragment:01135
fragment_chain:1"
6386. .14509
/note="assembly_fragment:01247
fragment_chain:2"
14610. .17060
/note="assembly_fragment:00977
fragment_chain:2"
17161. .21976
/note="assembly_fragment:01246
fragment_chain:2"
22077. .30708
/note="assembly_fragment:00221
fragment_chain:2"
30809. .46192
/note="assembly_fragment:01429
fragment_chain:3"
46293. .89106
/note="assembly_fragment:00855
fragment_chain:3"
89207. .97724
/note="assembly_fragment:00340"
97825. .103968
/note="assembly_fragment:00436"
104069. .106312
/note="assembly_fragment:00472"
106413. .110483
/note="assembly_fragment:00534"
110584. .117466
/note="assembly_fragment:00697"
117567. .127508
/note="assembly_fragment:00812"
127609. .130184
/note="assembly_fragment:01259
fragment_chain:4"
130285. .132386
/note="assembly_fragment:00064
fragment_chain:4"
132487. .141197
/note="assembly_fragment:00851
fragment_chain:4"
141298. .147891

```

BASE COUNT	a	c	g	t	others
ORIGIN	44319	26405	27064	48391	1712

Query Match	4.6%;	Score 55.4;	DB 2;	Length 147891;
Best Local Similarity	49.8%;	Pred. No. 0.05;		
Matches 140;	Conservative	0;	Mismatches 141;	Indels 0;
				Gaps 0;

QY	928	atgaacggaactcgttaagaatlcttlltgctaagcaagctattttaacagaagtltgtatctt	987
Dd	108584	AAGAATGAGACCTAGAAATTGGCTAGCACAGCAGGGTAACATAAGTCAAAAATAATTATATA	108525
QY	988	tttccagaaatctaggacaaggtcaacctgctgcggattaattractaggaattlctc	1047
Dd	108524	ATTCCTAGGTATATATTTTATGTGTGGCGTGTTGTAATAATAGGAATACGTTTAATTTCTTT	108465
QY	1048	ttccagtttagtcctgtatcttatttgataatcttcactatttgatcttgtatgatcttc	1107
Dd	108464	TTTCAGATTGTTTCATAGATGSTATATAGAAATGCTACTGATTTGTGAGCGTTGATTTTGT	108405
QY	1108	ttccctaaaaatttataaatttcctcaatctctgtaagtaattgaaatygatatattgtactt	1167
Dd	108404	ATCCGTGAAACTTTAATGATATTTATCATAGTCTCAATAGTTTTTTTGGAGACTATTACAGGTTT	108345

QY 1168 tctgtcaataatagaacacagacattgcgcaaaaaaaaaa 1208
| | | | | | | | | | | | | | | |
Db 108344 TTTTCAATATAAGATTCACATCATCTGCACAGATAATTAA 108304

RESULT	7
AC019074/c	
LOCUS	183892 bp DNA linear HTG 07-JUL-2000
DEFINITION	Homo sapiens chromosome X clone Rpl1-305M20, WORKING DRAFT
ACCESSION	AC019074
VERSION	AC019074.3 GI:7923976
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE	human.
ORGANISM	Homo sapiens

REFERENCE	AUTHORS	TITLE	JOURNAL	REFERENCE	AUTHORS	TITLE	JOURNAL
1	(bases 1 to 183892)	Waterston, R.H.	The sequence of Homo sapiens clone	2	(bases 1 to 183892)	Waterston, R.H.	Submitted (30-DEC-1999)
		Unpublished				Direct Submission	Genome Sequencing Center, Washington

COMMENT On May 18, 2000 this sequence version replaced gl:7022215.

```

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
----- Project Information -----
Center project name: H_NH0305M20
----- Summary Statistics -----
Sequencing vector: M13; 87%
Sequencing vector: plasmid; 13%
Chemistry: Dye-primer ET; 87% of reads
Chemistry: Dye-terminator Big Dye; 13% of reads
Assembly program: Phrap; version 0.990319
Consensus quality: 178146 bases at least Q40
Consensus quality: 179992 bases at least Q30
Consensus quality: 180917 bases at least Q20
Insert size: 189000; agarose-fp
Insert size: 182792; sum-of-contigs
Quality coverage: 4.59 in Q20 bases; agarose-fp
Quality coverage: 4.78 in Q20 bases; sum-of-contigs
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 12 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
1      1257: contig of 1257 bp in length
*      1258      1357: gap of unknown length
*      1358      3652: contig of 2295 bp in length
*      3653      3752: gap of unknown length
*      3753      10739: contig of 6987 bp in length
*      10740      10839: gap of unknown length
*      10840      18955: contig of 8116 bp in length
*      18956      19055: gap of unknown length
*      19056      29430: contig of 10375 bp in length
*      29431      29530: gap of unknown length
*      29531      45271: contig of 15741 bp in length
*      45272      45371: gap of unknown length
*      45372      59160: contig of 13789 bp in length
*      59161      59260: gap of unknown length
*      59261      78247: contig of 18987 bp in length
*      78248      78347: gap of unknown length

```


* 18558 19722: contig of 1165 bp in length
* 19723 19822: gap of unknown length
* 19823 21952: contig of 2130 bp in length
* 21953 22052: gap of unknown length
* 22053 23772: contig of 1720 bp in length
* 23773 23872: gap of unknown length
* 23873 25142: contig of 1270 bp in length
* 25143 25242: gap of unknown length
* 25243 26772: contig of 1530 bp in length
* 26773 26872: gap of unknown length
* 26873 28394: contig of 1522 bp in length
* 28395 28494: gap of unknown length
* 28495 29921: contig of 1427 bp in length
* 29922 30021: gap of unknown length
* 30022 31830: contig of 1809 bp in length
* 31831 31930: gap of unknown length
* 31931 33273: contig of 1343 bp in length
* 33274 33373: gap of unknown length
* 33374 35937: contig of 2564 bp in length
* 35938 36037: gap of unknown length
* 36038 38478: contig of 2441 bp in length
* 38479 38578: gap of unknown length
* 38579 41010: contig of 2432 bp in length
* 41011 41110: gap of unknown length
* 41111 43732: contig of 2622 bp in length
* 43733 43832: gap of unknown length
* 43833 45367: contig of 1535 bp in length
* 45368 45467: gap of unknown length
* 45468 47580: contig of 2113 bp in length
* 47581 47680: gap of unknown length
* 47681 49453: contig of 1773 bp in length
* 49454 49553: gap of unknown length
* 49554 51434: contig of 1881 bp in length
* 51435 51534: gap of unknown length
* 51535 55538: contig of 4004 bp in length
* 55539 55638: gap of unknown length
* 55639 59217: contig of 3579 bp in length
* 59218 59317: gap of unknown length
* 59318 63649: contig of 4332 bp in length
* 63650 63749: gap of unknown length
* 63750 68922: contig of 5173 bp in length
* 68923 69022: gap of unknown length
* 69023 73758: contig of 4736 bp in length
* 73759 73858: gap of unknown length
* 73859 79377: contig of 5519 bp in length
* 79378 79477: gap of unknown length
* 79478 86657: contig of 7180 bp in length
* 86658 86757: gap of unknown length
* 86758 95803: contig of 9046 bp in length
* 95804 95903: gap of unknown length
* 95904 107023: contig of 11120 bp in length
* 107024 107123: gap of unknown length
* 107124 119084: contig of 11961 bp in length
* 119085 119184: gap of unknown length
* 119185 133685: contig of 14501 bp in length
* 133686 133785: gap of unknown length
* 133786 144764: contig of 10979 bp in length
* 144765 144864: gap of unknown length
* 144865 156086: contig of 11222 bp in length
* 156087 156186: gap of unknown length
* 156187 170774: contig of 14588 bp in length
* 170775 170874: gap of unknown length
* 170875 192194: contig of 21320 bp in length
* 192195 192294: gap of unknown length
* 192295 216680: contig of 24386 bp in length.

FEATURES

Source

1. .216680
/organism="Homo sapiens"
/db_xref="taxon:9606"
/chromosome="17"
/clone="RP11-407E4"

BASE COUNT
ORIGIN

67337 a 38520 c 39310 g 66933 t 4580 others

Query Match 4.6%; Score 55.4; DB 2; Length 216680;
Best Local Similarity 49.8%; Pred. No. 0.049;
Matches 140; Conservative 0; Mismatches 141; Indels 0; Gaps 0;

QY 928 atgaacagaactgtaagaatttctgtagtaagaactatttaacagaagttgtgattt 987
DB 50223 AAGAAATAGGACCTAGAAATTTGCTAGACAGCAGGGTAACATATAGTCAAAAATATTATA 50282
QY 988 tttaacagaaatctagacaaggtcaacctgtctgctgcgattaactagatttctt 1047
DB 50283 ATTCTAGGTATATAATTTTATGTGTGGCTGTGTAATAGGATTAAGTTTAAATTTCTTT 50342
QY 1048 ttccagtttagtcctgtatttattgatatcttaccatttgattgtgatattt 1107
DB 50343 TTTCAGATTGTTCAATAGATGCTATATAGAAATGCTACTGATTGTGAGTGTGATTTCGT 50402
QY 1108 ttccctaaattttataatttccctaattcttgtaagtaattgaatgatattgtactt 1167
DB 50403 ATCCTGAACCTTAATGTATTATTATCAGTTCTAATAGTTTTTTGTGAGATATTCAGTTT 50462
QY 1168 tctgtcaataatagaacaagacattcgcaaaaaa 1208
DB 50463 TTTTCAATATATAGATCACATCATCTGCAACGATATTAA 50503

RESULT 9
AF274444/c
LOCUS
DEFINITION
AF274444 19209 bp DNA linear BCT 11-MAR-2001
L1 (rpl1), ribosomal protein L11 (rpl1), ribosomal protein
(rpl7/l12), beta subunit of RNA polymerase (rpoB), beta-prime
subunit of RNA polymerase (rpoC), ribosomal protein S12 (rps12),
ribosomal protein S7 (rps7), elongation factor G (fusa), elongation
factor Tu (tufA), ribosomal protein S10 (rps10), ribosomal protein
L3 (rpl3), ribosomal protein L4 (rpl4), ribosomal protein L2
(rpl2), ribosomal protein S3 (rps3), ribosomal protein L16
(rpl16), ribosomal protein S17 (rps17), ribosomal protein L14
(rpl14), ribosomal protein L5 (rpl5), ribosomal protein S14
(rps14), and ribosomal protein S8 (rps8) genes, complete cds; and
ribosomal protein L6 (rpl6) gene, partial cds.
AF274444
AF274444.1 GI:13272287

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Candidatus Carsonella ruddii.
Candidatus Carsonella ruddii.
Bacteria; Proteobacteria; gamma subdivision; Candidatus Carsonella.

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
1 (bases 1 to 19209)
Clark, M.A., Baumann, L., Thao, M.L., Moran, N.A. and Baumann, P.
Degenerative Minimalism in the Genome of a Psyllid Endosymbiont
J. Bacteriol. 183 (6), 1853-1861 (2001)
11222582

REFERENCE
AUTHORS
TITLE
JOURNAL
PUBMED
2 (bases 1 to 19209)
Clark, M.A., Baumann, L., Thao, M.L., Moran, N.A. and Baumann, P.
Direct Submission
Submitted (01-JUN-2000) Microbiology Section, University of
California Davis, One Shields Avenue, Davis, CA 95616-8665, USA
1. .19209
Location/Qualifiers

FEATURES

Source

/organism="Candidatus Carsonella ruddii"
/specific_host="Pachypsylla venusta"
/db_xref="taxon:114186"

gene
CDS

221. .649
/gene="rpl11"
221. .649
/gene="rpl11"
/codon_start=1
/transl_table=11
/product="ribosomal protein L11"
/protein_id="AKI7071.1"
/db_xref="GI:13272288"

```
gene
CDS
/translation="MKILFKLKLILKPNKATPSSGIGPILGQYGINLIDFCNKENDL
TKFICLEKINVKIAIFNDNSYQILINTMSLSDYIKEKKNTLSFSKKPGYINQLTNINLN
DVEISYFRNTYEKKKLNLSLRKMMGLTISMGYFNEKKYK"
630. .1205
/gene="rpl1"
630. .1205
/gene="rpl1"
/codon_start=1
/translation_table=11
/product="ribosomal protein L1"
/protein_id="AAK17072.1"
/db_xref="GI:13272289"
/translation="MKKSINDFLANLKKKKNFIESIDLNLILLKKKKYFNLSTNLF
YSYNEMKILFPLSNIOKIENNVYIGNFIDFLLKKVNSKVYTDSENLVLVKEKIGK
INXSIIYNCLXKNYEIEKNKFYKNTINLVLNKNNTLNKASTVYFYSMTIIRKNEFLI
FNLKKNFFFPNNIFIEKIYISSTMSKSFLLK"
1202. .1642
/gene="rpl10"
1202. .1642
/gene="rpl10"
/codon_start=1
/product="ribosomal protein L10"
/protein_id="AAK17073.1"
/db_xref="GI:13272290"
/translation="MNSYNIIRKNIKFNLNKYIILFYNEEKIKFELIDNKKKNT
KEFCSNKIIFEVKFNKSLYLHKNDLVYSLINNIDLIDIKPLLIINKNTQY
YDLIEKNSKLSKENLILKLILKNTIILKLNLIKNKYQYCK"
1620. .1961
/gene="rpl7/L12"
1620. .1961
/gene="rpl7/L12"
/codon_start=1
/product="ribosomal protein L7/L12"
/protein_id="AAK17074.1"
/db_xref="GI:13272291"
/translation="MENNIVNNVLDISKLNLVEITTLISNLEKKFNLYSNEKIAEK
NKFDIYLKEIGNKLNLIKTIKDITNMLNLESKSLVDNKNSLIKKNDKCEDEIVKK
ITLAGIVEIK"
1958. .5761
/gene="rpoB"
1958. .5761
/gene="rpoB"
/codon_start=1
/translation_table=11
/product="beta subunit of RNA polymerase"
/protein_id="AAK17075.1"
/db_xref="GI:13272292"
/translation="MINYLFKFRFCFSKKKFFHKCNLPYLLHNOIGSYNSFLSKYNS
LFSIKKVLKQYFPLICNNKNIFIKKLELLEPCNTEKYTKIRNLHSTLYMTISY
VINKNLNVYKKIFLIGNIPSMTKGNPIINGIDRILISQTKSYGIYFTEKKKKCI
IPLKGSWLEFIITNNFLIVFDKKINEINVELICLYNKKYFENEFKIKIKILRGK
KRIFLNNKKYFYKIYKICISIKILGKLFGKNYFLKNTFLKNIIDINKYNI
LNYPIYFFPIIFLKKIINLNYLIYIKNFELINKELIYKKGYIINFRLFYKKY
FYSLIGAKRIFKRLFLKKNKLTITYETIKILKFIKFNIONDNFDNLKLLNGCK
LSTIKEDFLFKYKFIYKMNNEFKYKDLDEIVNSDIITIGLDYFCNNELSQFLDQ
NMPLAEISHNRKISLISGIGIEKENGDIRDIHSHYCKVCPIDTPEGHNIGLNSL
AYLSKVNKYNFSTIYKQFIGKELGITFLDNFKDKFIYFNSTIETIYGEIFRSP
YFEARKTNYIYKFKFLNIDLIEICGDQISVGASLIPLSHNDANRCLMGSMQROAY
PLIDSENPVIGNELEIGLSNNTISDLNGVLYSDNYKIIKNNNFIKTYPLEKY
TRTNQNTILNOYTKVLKGFVKVGNITADSNSTKNGEISLGKNLVAFMSWGYNFE
SILLSSSILNNNSIHIYEFTVLYKNENGFEIYSNECGSNEKIKNKVKNGIKI
GEFVFSKDVIVGKMLPKKKRFSPEERLEKIVFSESNFNYEQPLVPKNIKGTIIAY
NDEKIFYFKNKIFKLLKEQNLNTCKNINNEFYETFNYYLRIKILLFNNKITIKKR
INSYNININNIIRKICFNKINFKLNIFKNIISNELKKNLFYVKKINFIKHDEEN
SIIRIKIKIAVKQOIAIGDKMSGHGKGVSNIIDYENMPYDKFGNKIDILLNPLG
VPSRMNVGQLEFVLGSLNLIKSFLKIKNLNKISYFKMLFIKIIFKCIYDKNINL
NIENNSVLKIFKNIKQNLNVCVHNFNENVNKVNNIITIGINKCELLLEDGITGK
RYLQLVNVGYIYFKMLNLVIDKIYSRSIGPYSIVTOOPLGGKSNLGGORLGEWEYA
LEAYGAFLKEMLTIKSDDILGRIELYKNIIGINDANSIPESFOVLMKEVQSLCF
DIKIL"
5758. .9636
gene
```

```
CDS
/gene="rpoC"
5758. .9636
/gene="rpoC"
/codon_start=1
/translation_table=11
/product="beta-prime subunit of RNA polymerase"
/protein_id="AAK17076.1"
/db_xref="GI:13272293"
/translation="MKSIFKVSIKMASSSKLISWSFGEVTHSNPTINEKNLKPEPN
LFCLEFSDLELCLCKKQCHSDKFLNLKHAHYGVGHIKLEFYAVHIWIKSSH
NISNTNLSTYKILEKTIINFKLVYIKSYDIKIKKENLFIIDKKIKNTIETLSAKALK
KLSDRELFIIDLIKIKIKKCNSEKLFSLKHINKLIFYYSGNKPSMVLKTPV
LPPKRPILALSEKKFASDINELKIVLERNYKLLKMTTYLDPKOLLINERISLOES
IVALEPNEKLNIPILITSSKRVLSFSSLIKGKGRFONLIGKRVDSGRSVITVEPN
LEVECKPIPIYALELEKPLIYELKKKITITISFIDDYKKNKOIILLKKNVKS
YFIMLNRAPTLHRMNFQSRKILTEDKTIKIHPLVCLSYNADFDGQMAIHLPTNNA
KVESNYLLSINNIISSPSNGNAIIPQDITMGIIYFTDFKKNKSISLINDIINEF
EFNNYNFQINIKFNNKKINTTLGRITINYLFKNEFDYFLNMVFKKILTYLVKYLFD
FNQLSKIIIIIEKLKIGFLISTYSGITISYFDLIEVKNFELINKVYFKLYNNNIES
IVELLNNEFVKIIRKIQINKTGAKTNNLFIMDSGRSGSMLOIKOLIAFGEFSKS
NGDITKDPIFDNLKNGLNKNYFISTFGARKGLDTSIKTANGSYLTRKLIDVSHDLV
IYKIDCKTKTGIKIFLNSYFTLVLYKNLGYRLLESVSLIKNKLKENLTIIDNKVLF
LLIKKKNITIFVRSVLEFCISNRGVCCFCYCIDISTNKIVLIGVSVGITSAGSIEGPT
OLMRTFHTGAVASYEYSDNLIISNSGFVKFKCKVCVINKFGEVLIYSLYGEMIFN
KLILEKYKFIYGTIKIFRNGFLIKNTKLINEDNNFYIYSENGIYFKEKNIVKN
YCLEDKKEYYKTLKNFKITLINSIKKNYEIFPKNENIVVYCYDIIPGEIIAKVSAT
ILKSSITIGLPRLSELEAPRPLKALLEISIDGCKIKFSLNLYITITISGFYKEY
TLCSEKRLXIYINNGDYKIGLIDSDGKPDLEINLISIDYLLNLYFNEINSIYEPONT
YVNCKHIELILKOMTKVKYIIFSGESSCOQGDILFLEDAINENFSTLVNSRRLSYKR
IVTGITKTSLESLSFESASFOETNKILVNSALIKNRTDYLGLKENYLVGRLLIPAGTG
LIKTFKRSKSSKNNLDFKRRKKFYINYF"
9638. .10000
/gene="rps12"
9638. .10000
/gene="rps12"
/codon_start=1
/translation_table=11
/product="ribosomal protein S12"
/protein_id="AAK17077.1"
/db_xref="GI:13272294"
/translation="MTLNQILKFKRKSYSKKRKPALLSSPQKKGICIKVYTTPKKP
NSALRKVCRAVKSLSKNKETAYIPGEGHNLQEHSNVLVGRGVKDLPGVKYHIIRNVYD
LSGVINRKTSRSKYGKKK"
10000. .10470
/gene="rps7"
10000. .10470
/gene="rps7"
/codon_start=1
/translation_table=11
/product="ribosomal protein S7"
/protein_id="AAK17078.1"
/db_xref="GI:13272295"
/translation="MSRKKRYEKTVIIIDNPKFGSYIIAKFISYIMNNGKKNLAQKIFY
YSIISIRLNKNPELILKILLYNVQPDFEIKRKKIGGSVKIPIKINLKRSLMFSMK
WIVKNSKLNENGYKKNLVGBLIDSYNNSLSTKQKDELNKIIDDNRKAYSNFKI"
10467. .12512
/gene="fusa"
10467. .12512
/gene="fusa"
/codon_start=1
/translation_table=11
/product="elongation factor G"
/protein_id="AAK17079.1"
/db_xref="GI:13272296"
/translation="MNDIKNIRNIGIIAHVDAGTTTTERILFFSGFSHKIGEVHTGN
TITDMKQEOERGITITTSASVTFEWMYNNSSINLIDTPGHVDFTEVERSLRVLDG
AVILICASSGIGQPOTEYVNOSEKENIPKILFVNKLDRIKAKYLSIENIKKKFECNT
LIININIGIENSFSGIDILINMKELIWNNSOLEIRNTNKNFDISNRYNILLLETLS
YDDITLEKYINSNFSIKDITESIRKLVILNKITPIACGSSLKNGIETLLDSIVNPLP
SPIDIGIKVNSNINSYVNIKSFKLALLFKVFNDYGLLSFIRIYSGKIEPGQIIFNN
SKNIKEKIFRIIRMFANSKKDLNIASAGDIVVLIGLKNSTFTGDTLSFDNEKVLLEKIN
IPLVIVSVSEPIVKNDEYKLNILNKFCCKEDPSLLEKINENTGELIISGMGELHLEI
IIDRINNEFNIKTKTSKPOYSYKESIKKTIIOEGKYIKQTGGRGOYGHVVLKIEPILI
```



```

CDS      5644. .6063
        /gene="MAL4P2.02"
        /note="MAL4P2.02, very hypothetical protein, len: 139 aa"
        /codon_start=1
        /product="very hypothetical protein, MAL4P2.02"
        /protein_id="CAB62843.1"
        /db_xref="GI:6562704"
        /translation="MYRNIIKYYFLTNMKRPTRYHYSRKONINLSNKKIDPYSDLKSS
        KYMKYKNVEYIPIKRILYKNASKKTNNNTYIIRMEK"
misc_feature 6787. .6792
        /note="alternative donor sequence (att/gtaaga) for
        alternative 5'-exon for MAL4P2.03"
gene      join(6794. .6857,6942. .7075,7272. .7383,7565. .7663,
        7792. .8064,8151. .8357,8454. .8588,8638. .9609,9740. .10018,
        10170. .10250,10340. .10378,10551. .10706,10776. .10787)
        /gene="MAL4P2.03"
CDS      join(6794. .6857,6942. .7075,7272. .7383,7565. .7663,
        7792. .8064,8151. .8357,8454. .8588,8638. .9609,9740. .10018,
        10170. .10250,10340. .10378,10551. .10706,10776. .10787)
        /gene="MAL4P2.03"
        /note="MAL4P2.03, hypothetical protein, len: 853 aa, 13
        exons but some splice predictions are tentative and
        alternatives exist"
        /codon_start=1
        /product="hypothetical protein"
        /protein_id="CAB62844.1"
        /db_xref="GI:6562705"
        /db_xref="SPTREMBL:Q9U0L9"
        /translation="MNVEDRYKLKEKYEKIEQNDILKKAILEYKDKLELEKKNDI
        LSNKNQLOKNLTQLTNSLEEQKKNSGWTNMLLTNSRENHKSVAFELEMKIKE
        NEELHKNIEDLHOEKSYLKELELLKDHIEKLENEKELTIGTLNDVNSSHEQIKAD
        EYELIKKLIENOEKESQIKKESLDYTRNISTMLHNKCPKPMINFDQIPFYQA
        YDREGSYIOKEILOENERANYFLSLQKVYVYLKCEKLEFLFOKNLDSMYNKKME
        KLEEDDKYSKMIDLKNCIDQKVLIEIIDTAEEKFIYAEKEDKELMEVFLNTYNI
        OMATHKNYFIHIIKKNIRELFEKVNVYLCEIEYMPSPFINNKVFLRNVIQOFRLKE
        LVKIIINITLIIISICPYNNERRIILENLRKNRIKKENKIIINHEKRNDEEKTNIS
        SFDMHITNIYDDDKLEFYNKOEYENKSLYNNFEIVEKKIPFINFOEKEIMLSOFND
        ETNYQTKNNHIIILFIKKIQRILOIYIPQSLKLSYISYRLCEINRGNFYINNSPLI
        TNLIDKTNEFIKSIENFNFIHIIYITLSYNTHTPYLQOHMYINNINOTFKA
        KVHTYDKLOKNLKDLOPKKDNIOIOLMKLKSKKENKKLKTTLNPNRNIILNQELSL
        NNQVLYNSIDNKREDKOYSDNNNIIOTYAEKIYVNFISHTDEGKNNRIKELQIEAYI
        SSCRIINNLEIIONNKHILELONLFNIKENIKNLNTQIDITYKEEVNIIHKKYEEO
        MNTLEDLIVSLEKQINKINSEKNVNFIFIMCSICSNKNNIGRKYNPICNMLYGLF"
misc_feature 6858. .6863
        /gene="MAL4P2.03"
        /note="predicted splice donor sequence, atg/gtaata, for
        exon 1 of MAL4P2.03"
misc_feature 6932. .6941
        /gene="MAL4P2.03"
        /note="predicted splice acceptor sequence for exon 2 of
        MAL4P2.03"
misc_feature 7076. .7081
        /gene="MAL4P2.03"
        /note="predicted splice donor sequence, caa/gtaaat, for
        exon 2 of MAL4P2.03"
misc_feature 7262. .7271
        /gene="MAL4P2.03"
        /note="predicted splice acceptor sequence for exon 3 of
        MAL4P2.03"
misc_feature 7384. .7389
        /gene="MAL4P2.03"
        /note="potential splice donor sequence, atg/gtaaaa, for
        exon 3 of MAL4P2.03"
misc_feature 7556. .7565
        /gene="MAL4P2.03"
        /note="predicted splice acceptor sequence for exon 4 of
        MAL4P2.03"
misc_feature 7664. .7669
        /gene="MAL4P2.03"
        /note="predicted splice donor sequence, aac/gtttgt, for
        exon 4 of MAL4P2.03"

```

```

misc_feature      7782..7791
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 5 of
MAL4P2.03"
                  8065..8070
                  /gene="MAL4P2.03"
                  /note="potential splice donor sequence, caa/gtaaaa, for
exon 5 of MAL4P2.03"
                  8141..8150
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 6 of
MAL4P2.03"
                  8358..8363
                  /gene="MAL4P2.03"
                  /note="potential splice donor sequence, cag/gtagga, for
exon 6 of MAL4P2.03"
                  8443..8453
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 7 of
MAL4P2.03"
                  8589..8594
                  /gene="MAL4P2.03"
                  /note="predicted splice donor sequence, tat/gtataat, for
exon 7 of MAL4P2.03"
                  8628..8637
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 8 of
MAL4P2.03"
                  9610..9615
                  /gene="MAL4P2.03"
                  /note="predicted splice donor sequence, acc/gtaaat, for
exon 8 of MAL4P2.03"
                  9645..9650
                  /gene="MAL4P2.03"
                  /note="alternative splice donor sequence, acc/gtaaat, for
exon 8 of MAL4P2.03"
                  9730..9739
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 9 f
MAL4P2.03"
                  9747..9756
                  /gene="MAL4P2.03"
                  /note="alternative splice acceptor sequence for exon 9 of
MAL4P2.03"
                  10019..10024
                  /gene="MAL4P2.03"
                  /note="predicted splice donor sequence, cac/gtaaga, for
exon 9 of MAL4P2.03"
                  10160..10169
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 10 of
MAL4P2.03"
                  10251..10256
                  /gene="MAL4P2.03"
                  /note="predicted splice donor sequence, aag/gtaata, for
exon 10 of MAL4P2.03"
                  10330..10339
                  /gene="MAL4P2.03"
                  /note="predicted splice acceptor sequence for exon 11 of
MAL4P2.03"

Query Match      4.4%; Score 53.8; DB 3; Length 234112;
Best Local Similarity 57.4%; Pred. No. 0.11;
Matches 97; Conservative 0; Mismatches 72; Indels 0; Gaps 0

QY 1042 ttctcttcaggttagtcctgtatttatttgatattcttaacctattgattgtatg 1101
||||| 111 111 111 111 111 111 111 111 111 111 111 111 111 111 111
TTTTTTTTCATTTATATCATATTTTATGAGTTATTTTATTTTATTTTATTTTATTTTATTTT
Db 110827

QY 1102 atttttcccttaaaatttataatttccctaattcttgtaagtaattgaaatgatat 1161
||||| 1 111 111111111111111 111 1 111 111 111 111 111 111 111
TTTTTGTATATACATTTTATATTTTCATATATTTGATATATTAAGAATAATACATTTTAAATC
Db 110887

```


GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2002, 20:27:32 ; Search time 194.53 Seconds

(without alignments)
10679.414 Million cell updates/sec

Title: US-09-805-919-3

Perfect score: 1210

Sequence: 1 tatattcaattgaaacatg.....ttcgcaaaaaaaaaaaaaaa 1210

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1736436 seqs, 858457221 residues

Total number of hits satisfying chosen parameters: 3472872

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database: X_Geneseq_032802.*

1: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1980.DAT:*
2: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1981.DAT:*
3: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1982.DAT:*
4: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1983.DAT:*
5: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1984.DAT:*
6: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1985.DAT:*
7: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1986.DAT:*
8: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1987.DAT:*
9: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1988.DAT:*
10: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1989.DAT:*
11: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1990.DAT:*
12: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1991.DAT:*
13: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:*
14: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:*
15: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1994.DAT:*
16: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:*
17: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:*
18: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1997.DAT:*
19: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1998.DAT:*
20: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:*
21: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:*
22: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:*
23: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:*
24: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1210	100.0	1210	24	ABA01665	Nicotiana tabacum
2	927	76.6	927	24	ABA01664	Nicotiana tabacum
3	62.2	5.1	612	22	AAH71471	Human cervical can
4	59.6	4.9	622	22	AAI20340	Human breast can
5	54.4	4.5	411	22	AAI88595	Human polynucleoti
6	54	4.5	11662	24	ABL33900	Human immune syste
7	53.4	4.4	8952	22	AAS46446	Tumour suppressor
8	53.4	4.4	17869	24	ABL32105	Human immune syste
9	53	4.4	198	22	AAS07727	Cervical cancer pr

C	10	52.8	4.4	337	22	AAI1578	Human breast cance
	11	52.6	4.3	394	22	AAI10020	Human breast cance
	12	52.4	4.3	425	22	AAS60450	Human cancer agent
	13	52.4	4.3	426	22	AAI11441	Human breast cance
	14	52.4	4.3	805	24	ABL32513	Human immune syste
	15	52.2	4.3	12592	24	AAS61101	Human gene regulat
	16	52.2	4.3	14950	24	ABL33257	Human immune syste
	17	52	4.3	393	22	AAI11189	Human breast cance
	18	51.6	4.3	8342	24	ABL32501	Human immune syste
	19	51.4	4.2	836	22	AAI94700	Human neuroblastom
C	20	51.2	4.2	377	22	AAI86808	Human polynucleoti
	21	51.2	4.2	56153	22	AAS46793	Tumour suppressor
	22	50.8	4.2	5449	24	ABL33641	Human immune syste
	23	50.8	4.2	7657	22	AAS45477	Chemically pretrea
	24	50.8	4.2	7657	24	ABL34022	Human immune syste
C	25	50.6	4.2	18624	24	ABL33703	Human neuroblastom
	26	50.4	4.2	875	22	AAI95044	Human immune syste
	27	50.4	4.2	9021	24	ABL34232	Human immune syste
	28	50.2	4.1	621	22	AAH71551	Human cervical can
C	29	50	4.1	1164	22	AAF98699	Human ovarian can
	30	49.8	4.1	278	22	AAH69999	Human cervical can
	31	49.8	4.1	5134	22	AAS45460	Chemically pretrea
	32	49.8	4.1	5134	24	ABL33819	Human immune syste
	33	49.8	4.1	7763	24	ABL34437	Human immune syste
	34	49.4	4.1	6044	24	ABL32840	Human immune syste
	35	49.4	4.1	6291	22	AAS46452	Tumour suppressor
	36	49.4	4.1	9636	15	AAQ67190	P. falciparum tran
C	37	49.2	4.1	250	22	AAS29153	cDNA encoding for
	38	49.2	4.1	545	22	AAH70126	Human cervical can
C	39	49.2	4.1	1336	22	AAS29081	cDNA encoding for
	40	49.2	4.1	14919	22	AAS46506	Tumour suppressor
	41	49	4.0	396	22	AAF94842	Human ovarian can
	42	49	4.0	6436	24	ABL32681	Human immune syste
	43	49	4.0	6478	24	ABL32498	Human immune syste
	44	49	4.0	17389	24	ABL33414	Human immune syste
	45	48.8	4.0	6112	24	ABL32663	Human immune syste

ALIGNMENTS

RESULT	1
ID	ABA01665 standard; DNA; 1210 BP.
AC	ABA01665;
XX	
DT	06-FEB-2002 (first entry)
DE	Nicotiana tabacum cv. Xanthinc stress induced gene SEQ ID NO:3.
XX	
KW	Nicotiana tabacum cv. Xanthinc; stress induced protein; plant;
KW	stress induced gene; transgenic plant; resistance; ds.
OS	Nicotiana tabacum.
XX	
FH	Key
FT	CDS
FT	Location/Qualifiers
FT	18..943
FT	/*tag= a
FT	/product= "stress induced protein"
FT	/trans_except= (pos:846..848,aa:Xaa)
PN	JP2001252084-A.
XX	
PD	18-SEP-2001.
XX	
PF	15-MAR-2000; 2000JP-0071655.
XX	
PR	15-MAR-2000; 2000JP-0071655.
XX	
PA	(NARA-) NARA SENTAN KAGAKU GIYUTSU DAIGAKUIN DAI.
XX	
DR	WPI; 2002-029660/04.

DR P-PSDB; AAG68236.

XX New polypeptide for creating plant resistance to injury, comprises a
PT gene, the expression of which is induced by stress
XX
XX
PS Claim 4; Page 8; 11pp; Japanese.

XX The present sequence represents a gene isolated from tobacco, in which
CC the expression is induced by stress, and so encodes a stress induced
CC protein (I). The present invention also describes: (1) providing
CC resistance against stress by introducing the above gene to a plant; and
CC (2) a transformed plant which obtained resistance against stress by
CC introducing the above gene to the plant. The gene can be used for
CC creating plants resistant against injury, osmotic pressure, salt, or
CC low temperature-induced stress.

XX Sequence 1210 BP; 347 A; 217 C; 251 G; 395 T; 0 other;

Query Match 100.0%; Score 1210; DB 24; Length 1210;
Best Local Similarity 100.0%; Pred. No. 1.3e-268;
Matches 1210; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 tataatcaattgaaaacatgtttgacaagaggcgtctttcgctgtgtttgttacttg 60
Db 1 tataatcaattgaaaacatgtttgacaagaggcgtctttcgctgtgtttgttacttg 60
QY 61 tgcacactcataaagcagttctaaagcgagatatttccaatgtgtccctcttcctgcg 120
Db 61 tgcacactcataaagcagttctaaagcgagatatttccaatgtgtccctcttcctgcg 120
QY 121 gtgatatcaataaataattccctccgactgagagactgatacccgagcatgtgtgtagac 180
Db 121 gtgatatcaataaataattccctccgactgagagactgatacccgagcatgtgtgtagac 180
QY 181 gcgatatgagctcgatgtccagaacaaccaaaccgtgttcaattacaatccagaattt 240
Db 181 gcgatatgagctcgatgtccagaacaaccaaaccgtgttcaattacaatccagaattt 240
QY 241 tcgacgtacaaggaattaactacaagaagctactcaataagctacttgatccctgcctaa 300
Db 241 tcgacgtacaaggaattaactacaagaagctactcaataagctacttgatccctgcctaa 300
QY 301 atgatacagagagaaaattgacagttttccaatcaacggcgcaagttatgatgcacatga 360
Db 301 atgatacagagagaaaattgacagttttccaatcaacggcgcaagttatgatgcacatga 360
QY 361 ctgaccaaactcttgaatgggtcgtgttaacaatgatatacaactatgtccaactgtctag 420
Db 361 ctgaccaaactcttgaatgggtcgtgttaacaatgatatacaactatgtccaactgtctag 420
QY 421 ctccatcaatcgttcacagtatattcctacaagttttttagcacaanaattcaacgggtt 480
Db 421 ctccatcaatcgttcacagtatattcctacaagttttttagcacaanaattcaacgggtt 480
QY 481 ttagctacctgttcataagaagaataattgcaagcttcggatttggctggcggtctaggg 540
Db 481 ttagctacctgttcataagaagaataattgcaagcttcggatttggctggcggtctaggg 540
QY 541 ttgaaactgttgcatgttcctctgtctccaggcatttcatcaacaagtcgltctaagttat 600
Db 541 ttgaaactgttgcatgttcctctgtctccaggcatttcatcaacaagtcgltctaagttat 600
QY 601 caagcacacatcaaggcctgtgttatgggtttgagctttcttggaaagcgttaactgttat 660
Db 601 caagcacacatcaaggcctgtgttatgggtttgagctttcttggaaagcgttaactgttat 660
QY 661 gtagaacttgcgaccggagtcgtgggggtgagtgcaactattgagaanaacagcgacagag 720
Db 661 gtagaacttgcgaccggagtcgtgggggtgagtgcaactattgagaanaacagcgacagag 720
QY 721 ctactgtcgttattgtgtgcaaaaggacatcacglttcgaagcttacggttcgatgca 780
Db 721 ctactgtcgttattgtgtgcaaaaggacatcacglttcgaagcttacggttcgatgca 780

Db 721 ctactgtcgttattgtgtgcaaaaggacatcacglttcgaagcttacggttcgatgca 780
QY 781 aagtcgagtaactatctgtttatgtatgtcttcttggcggtataggaatagtgtagttt 840
Db 781 aagtcgagtaactatctgtttatgtatgtcttcttggcggtataggaatagtgtagttt 840
QY 841 tggcgctaagatttccactaggaattccaacttgcacgcagcgtgtgtgacgtgca 900
Db 841 tggcgctaagatttccactaggaattccaacttgcacgcagcgtgtgtgacgtgca 900
QY 901 aaagacggaatttgcatatacctcccgatgaaacagaactgttaagaatttgcagtaaa 960
Db 901 aaagacggaatttgcatatacctcccgatgaaacagaactgttaagaatttgcagtaaa 960
QY 961 gctattttaacagaagtttgtgtattttttcagaaaatctaggaacaaggtcaacctgtg 1020
Db 961 gctattttaacagaagtttgtgtattttttcagaaaatctaggaacaaggtcaacctgtg 1020
QY 1021 ctggcgatttaactaggaatttcttccaggtttagtcctgtatttatttgatattc 1080
Db 1021 ctggcgatttaactaggaatttcttccaggtttagtcctgtatttatttgatattc 1080
QY 1081 ttacctattgtatgtgtatgatttttcccttaaaaatttataaatttccctaattcttg 1140
Db 1081 ttacctattgtatgtgtatgatttttcccttaaaaatttataaatttccctaattcttg 1140
QY 1141 taagtaattgaatgatatattgtacttctgtcaataatagacaagaacattcgcaaaaa 1200
Db 1141 taagtaattgaatgatatattgtacttctgtcaataatagacaagaacattcgcaaaaa 1200
QY 1201 aaaaaaaaaa 1210
Db 1201 aaaaaaaaaa 1210

RESULT 2
ABA01664
ID ABA01664 standard; DNA; 927 BP.

XX ABA01664;
AC
XX 06-FEB-2002 (first entry)
DT
XX
DE Nicotiana tabacum cv. Xanthine stress induced gene SEQ ID NO:2.
XX
KW Nicotiana tabacum cv. Xanthine; stress induced protein; plant;
KW stress induced gene; transgenic plant; resistance; ds.
XX Nicotiana tabacum.

XX
FH Key Location/Qualifiers
FT CDS 1..927
FT /*tag= a
FT /product= "stress induced protein"
FT /trans_except= (pos:828..830,aa:Xaa)

PN JP2001252084-A.

PD 18-SEP-2001.

XX 15-MAR-2000; 2000JP-0071655.

PR 15-MAR-2000; 2000JP-0071655.

XX (NARA-) NARA SENTAN KAGAKU GIJUTSU DAIGAKUIN DAI.

DR WPI; 2002-029660/04.

DR P-PSDB; AAG68236.

XX New polypeptide for creating plant resistance to injury, comprises a
PT gene, the expression of which is induced by stress
XX
XX
PS Claim 3; Page 7-8; 11pp; Japanese.

Db 342 TAAATGTTTTTTTTTATTTTAAATTAGTTTTTTTTTTTTTTTTTTTATT 283
QY 967 ttaacagaagttgtgtatcttttcagaaaaatcaggacaaggtcaacctgtgtgcg 1026
Db 282 TTAAGTTTATTTTATTTTATTTTATTTTAAATAATTTTGGTTAGTTGAATAAT 223
QY 1027 attaatcactaggattttcttccagtttagtcctgtatttattgataattcctaac 1086
Db 222 GGGTATGGGTGATTTTATTTTATTTTATTTTGTGTTTTTTTTTTTTTTTTTTT 163
QY 1087 attgattgtgtatgattttttcccttaaaatttataatttccctaattctgtaagta 1146
Db 162 TTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 103
QY 1147 attgaatgataattgttacttctgtcaataatagacaagacattcgcaaaaaaaaa 1206
Db 102 ATTAAAAAAATTAAAAAAATTTAAAAAAATTTAAAAAAATTTAAAAAAATTTAAAA 43
QY 1207 aaaa 1210
Db 42 AAAA 39

RESULT 4

AAI20340

ID AAL20340 standard; cDNA; 622 BP.

AC AAL20340;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 12797.

KW Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.

PN WO200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077.

PR 14-MAR-2000; 2000US-0189167.

PR 24-MAR-2000; 2000US-0192099.

PR 29-MAR-2000; 2000US-0193480.

PR 15-MAY-2000; 2000US-0205230.

PR 09-JUN-2000; 2000US-0211315.

PR 25-JUL-2000; 2000US-0220534.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI; 2001-451856/48.

PS New peptide useful as a marker for the diagnosis of breast cancer

PS Claim 1; Page 2261-2262; 3695pp; English.

CC The invention relates to human breast cancer expressed polynucleotides (AAI07544-AAI26789) and methods of assessing whether a patient is afflicted with breast cancer by examining the correlation between the expression of certain markers and the cancerous state of breast cells. The polynucleotides and encoded polypeptides are potential markers for detecting, diagnosing, monitoring, characterising treating and potentially preventing breast cancer. The polynucleotides and encoded CC polypeptides are also useful for isolating compounds with cytostatic activity.

SO Sequence 622 BP; 189 A; 154 C; 91 G; 188 T; 0 other;

Query Match 4.9%; Score 59.6; DB 22; Length 622;
Best Local Similarity 59.4%; Pred. No. 0.00024;
Matches 101; Conservative 0; Mismatches 69; Indels 0; Gaps 0;

QY 1041 ttttcttccagtttagtcctgtatttatttgatatcttacctatttgatgtgtat 1100
Db 46 ttt 105
QY 1101 gatttttttccctaaattttataattttccctaattcttgaagtaattgaatgatat 1160
Db 106 ttt 165
QY 1161 tgtacttctgtcaataatagacaagacattcgcaaaaaaaaaaaaaa 1210
Db 166 tttttgggggaaaaaaaaaaaaaaaaaaccacaaaaaaaaaaaaa 215

RESULT 5

AAI88595/C

ID AAI88595 standard; cDNA; 411 BP.

AC AAI88595;

DT 06-NOV-2001 (first entry)

DE Human polynucleotide SEQ ID NO 8655.

KW Human; cytokine; cell proliferation; cell differentiation; gene therapy;

KW vaccine; peptide therapy; stem cell growth factor; haematopoiesis;

KW tissue growth factor; immunomodulatory; cancer; leukaemia;

OS Homo sapiens.

PN WO200164835-A2.

PD 07-SEP-2001.

PF 26-FEB-2001; 2001WO-US04927.

PR 28-FEB-2000; 2000US-0515126.

PR 18-MAY-2000; 2000US-0577409.

PA (HYSE-) HYSEQ INC.

PI Tang YT, Liu C, Drmanac RT;

DR WPI; 2001-514838/56.

DR P-PSDB; AAC08664.

PT Isolated nucleic acids and polypeptides, useful for preventing

PT diagnosing and treating e.g. leukaemia, inflammation and immune

PT disorders

PS Claim 1; SEQ ID NO 8655; 1399pp + Sequence Listing; English.

CC The invention relates to human polynucleotides (AAI79941-AAI93841) and the encoded proteins (AAO0010-AAO13910) that exhibit activity elating to cytokine, cell proliferation or cell differentiation or which may induce CC production of other cytokines in other cell populations. The CC polynucleotides and polypeptides are useful in gene therapy, vaccines or peptide therapy. The polypeptides have various cytokine-like activities, e.g. stem cell growth factor activity, immunomodulatory activity and activity/inhibin activity and may be useful in the diagnosis and/or treatment of cancer, leukaemia, nervous system disorders, arthritis and inflammation.

CC Note: The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.

SO Sequence 411 BP; 214 A; 24 C; 64 G; 107 T; 2 other;

Query Match	4.5%;	Score 54.4;	DB 22;	Length 411;
Best Local Similarity	51.2%;	Pred. No. 0.0034;		
Matches 127;	Conservative	0;	Mismatches 121;	Indels 0;
				Gaps 0;

[illegible]

RESULT	6
ABL33900	
ID	ABL33900 standard; DNA; 11662 BP.
XX	
AC	ABL33900;
XX	
DT	26-MAR-2002 (first entry)
XX	
DE	Human immune system associated gene SEQ ID NO: 1873.
XX	
KW	Human; immune system disease; cytosine methylation; antiasthmatic;
KW	antiarteriosclerotic; antianaemic; cytotatic; nootropic;
KW	neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW	antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW	antiflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW	acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW	neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW	gene; ds.
XX	
OS	Homo sapiens.
XX	
PN	WO200200928-A2.
XX	
PD	03-JAN-2002.
XX	
PF	02-JUL-2001; 2001WO-EP07537.
XX	
PR	30-JUN-2000; 2000DE-1032529.
PR	01-SEP-2000; 2000DE-1043826.
XX	
PA	(EPIG-) EPIGENOMICS AG.
XX	
PI	Olek A, Plepenbrock C, Berlin K;
XX	
DR	WPI; 2002-130909/17.
XX	
PT	Nucleic acid comprising fragment of chemically modified gene, useful
PT	for diagnosis and treatment of diseases associated with abnormal
PT	cytosine methylation -
XX	
PS	Claim 1; SEQ ID NO 1873; 32pp + Sequence Listing; German.
XX	
CC	The present invention provides a number of human immune system associated
CC	genes which are modified by the methylation of cytosines. The sequences
CC	can be used in the diagnosis and treatment of immune system disorders,
CC	including eye diseases such as retinopathy, neovascular glaucoma and

CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 11662 BP; 3110 A; 120 C; 2632 G; 5800 T; 0 other;

Query Match	4.5%;	Score 54;	DB 24;	Length 11662;
Best Local Similarity	57.8%;	Pred. No. 0.0098;		
Matches	96;	Conservative	0;	Mismatches 70;
			Indels	0;
			Gaps	0;

QY 1027 atcaattactaggaattttcttccagtttagtcgtgattttatgtgatattcctaact 1086
||| ||| | | | | | | | | | | | | |
Db 92 attattattagtgataattaagttagtatattaagagtattttttcgatgttttttt 151

QY 1087 attgtagtgtatgatttttttccttaaaattcataallttccaattctgtgaagta 1146
||| ||| | | | | | | | | | | | | |
Db 152 ttgcgattacatttttgttttttttataaatlttgtattttaattgtgatttggtaaca 211

QY 1147 attgaatggatatatttgtaaccttctgtcaataataagaacagacatt 1192
||| ||| | | | | | | | | | | | | |
Db 212 atatagatcatatttatttgttttcoggtatttaataataagaattt 257

RESULT	7	
ID	AAS46446	standard; DNA: 8952 BP.
AC	AAS46446;	
DT	18-DEC-2001	(first entry)
DE	Tumour suppressor gene derived chemically modified sequence #168.	
KW	Human; tumour suppressor gene; oncogene; antitumour; cytostatic;	
KW	cancer; tumour; CpG dinucleotide; single-nucleotide polymorphism; SNP;	
KW	cytosine methylation; ds.	
OS	Homo sapiens.	
PN	WO200168912-A2.	
PD	20-SEP-2001.	
PF	15-MAR-2001; 2001WO-EP02955.	
PR	15-MAR-2000; 2000DE-1013847.	
PR	06-APR-2000; 2000DE-1019058.	
PR	07-APR-2000; 2000DE-1019173.	
PR	30-JUN-2000; 2000DE-1032529.	
PR	01-SEP-2000; 2000DE-1043826.	
PA	(EPIG-) EPIGENOMICS AG.	
PI	Olek A, Piepenbrock C, Berlin K;	
DR	WPI; 2001-602752/68.	
PT	Fragments of chemically modified genes associated with tumour suppressor	
PT	genes and oncogenes, useful in designing primers and probes for	
PT	analysing diseases associated with cytosine methylation state e.g.	
PT	cancer	
PS	Claim 1; SEQ ID No 168; 27pp; English.	
CC	The invention relates to a nucleic acid comprising a sequence of 18	
CC	bases, of a segment of chemically pretreated DNA (CP DNA) e.g. with	
CC	bisulphite, of genes associated with tumour suppression and	
CC	oncogenes having a sequence taken from 536 (actually 533 since	
CC	numbers 408, 458 and 500 are missing from the sequence listing) sequences	
CC	(Ss) and sequences complementary to (Ss). The nucleic acid may be a	
CC	peptide nucleic acid-oligomer (PNA) of at least 9 nucleotides and may	

CC form part of a set of probes for detecting the cytosine methylation state
 CC and/or single nucleotide polymorphisms and also to be used in an
 CC array for analysing diseases associated with CpG dinucleotides e.g.
 CC cancers and tumours. The probes can also be used in a method for
 CC ascertaining genetic and/or epigenetic parameters for the diagnosis
 CC and/or therapy of existing diseases or the predisposition to specific
 CC diseases, by analysing cytosine methylations. The parameters may be
 CC compared to another set of genetic and/or epigenetic parameters, the
 CC differences serving as basis for diagnosis and/or prognosis events which
 CC are disadvantageous to patients. The present sequence is one of the
 CC 533 genomic sequences derived from tumour suppressor genes and
 CC oncogenes. Sequences with even numbered Seq ID numbers are the
 CC complementary sequence of the corresponding odd numbered sequence (e.g.
 CC ID 2 and ID1, ID 536 and ID 535, except for those whose partner sequence
 CC is missing).
 CC Note: The sequence data for this patent did not form part
 CC of the printed specification, but was obtained in electronic
 CC format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 8952 BP; 2404 A; 162 C; 1905 G; 4481 T; 0 other;

Query Match	4.4%;	Score 53.4;	DB 22;	Length 8952;
Best Local Similarity	52.5%;	Pred. NO. 0.013;		
Matches 117;	Conservative	0;	Mismatches 106;	Indels 0;
				Gaps 0;
QY 964 atttaacagaagttgtgtatcttttcagaaatctaggacaagtcacacctgtgctg	1023			
Db 8529 aatttatctttgttttgggacutttgttagagaaagtataaatagtgtaaaggttacgg	8588			
QY 1024 gcgattaaactactagatttttcccttcocagttttagtcctgtattttattgataattccta	1083			
Db 8589 tataattatttattatgttttttcaatttagttatgattgtttttatataattgtagtttttttatt	8648			
QY 1084 cctatttgattgtgatgatttttttcccttaaaattttataaattttccctaattcttgttaa	1143			
Db 8649 ttttttttattgattataaagtagttgtttttaaatttttttattttttttaagtttgyta	8708			
QY 1144 gtaattgaatggataattgttactttctgtcaataatagaacaa	1186			
Db 8709 tataattttgtttattatttattttgttttttaatgaattttataa	8751			

RESULT	ID	ABL32105	standard; DNA; 17869 BP.
XX	ID	ABL32105	standard; DNA; 17869 BP.
XX	AC	ABL32105;	
XX	DT	26-MAR-2002	(first entry)
XX	DE	Human immune system associated gene SEQ ID NO: 78.	
KW		Human; immune system disease; cytosine methylation; antiasthmatic;	
KW		antiarteriosclerotic; antihaemic; cyostatic; nootropic;	
KW		neuroprotective; anti-HIV; anticonvulsant; ophthalmological;	
KW		antirheumatic; antiarthritic; antidiabetic; antipsoriatic;	
KW		antinflamatory; cancer; eye disease; arteriosclerosis; anaemia;	
KW		acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;	
KW		neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;	
KW		gene; ds.	
OS		Homo sapiens.	
XX	PN	WC0200200928-A2.	
XX	PD	03-JAN-2002.	
XX	PF	02-JUL-2001; 2001WO-EP07537.	
XX	PR	30-JUN-2000; 2000DE-1032529.	
PR		01-SEP-2000; 2000DE-1043826.	

XX (EPiG-) EPIGENOMICS AG.
PA
XX
XX Olek A, Piepenbrock C, Berlin K;
PI
XX
DR WPI; 2002-130909/17.
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
PS Claim 1; SEQ ID NO 78; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
XX
XX Sequence 17869 BP; 5366 A; 158 C; 3365 G; 8978 T; 2 other;

	Query Match	4.48;	Score 53.4;	DB 24;	Length 17869;	
	Best Local Similarity	54.3%;	Pred. No. 0.015;			
	Matches 108; Conservative	0;	Mismatches 91;	Indels 0;	Gaps 0;	
QY	963 tatttaacagaagttgtgatatcttccagaaatctagacaagtcaacctgtgct	1022				
Dd	1971 tgttttatagtttagtatcttcctttaagggatgttggggaagaattaatcaatt	2030				
QY	1023 ggcgattaactactaggaatttcctcaccagtttagtcctgcatlltatttgataattc	1082				
Dd	2031 tgttcgaaattctaqtllatttgtlataltttaagttctaagtcgtgttttattttaagaagtttg	2090				
QY	1083 acctatttgattgtgtatgaatttttcccttaaattttatataatttcctaattcttgta	1142				
Dd	2091 aaqttglttaatgtataatgttttctaagtttatttttatacataattttttttttgtt	2150				
QY	1143 agtaatgaatggaataatt	1161				
Dd	2151 aatgatatataataattttt	2169				

RESULT	9
AAS07727	
ID	AAS07727 standard; DNA; 198 BP.
XX	
AC	AAS07727;
XX	
DT	23-OCT-2001 (first entry)
XX	
DE	Cervical cancer pre-malignant condition DNA marker #26.
XX	
KW	Cervical cancer; pre-malignant condition marker; tumourigenesis; CIN; ds
KW	cervical intraepithelial neoplasia; squamous intraepithelial lesions;
KW	polymerase chain reaction; PCR; enzyme linked immunosorbent assay; ELISA
KW	anticancer therapy; carcinogen; antisense inhibition.
XX	
OS	Homo sapiens.
XX	
PN	WO200142792-A2.
XX	
PD	14-JUN-2001.
XX	
PF	08-DEC-2000; 2000WO-US33311.
XX	
PR	08-DEC-1999; 99US-0169811.
PR	21-DEC-1999; 99US-0171330.
PR	14-MAR-2000; 2000US-0189113.
PR	31-MAR-2000; 2000US-0193943.

PR 12-MAY-2000; 2000US-0203772.
PR 09-JUN-2000; 2000US-0210820;
PR 21-JUL-2000; 2000US-0220113.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
XX
PI Schlegel R, Deeds J, Berger A, Zhao X;
XX
DR WPI: 2001-367889/38.
XX
XX Cervical cancer protein markers useful for the diagnosis, prevention
PT and treatment of cervical cancers, especially cervical intraepithelial
PT neoplasia or squamous intraepithelial lesions -
XX
PS 20; Page 407; 436pp; English.

CC The sequence represents a cervical cancer pre-malignant condition marker.
 CC These markers encode proteins that are over expressed during
 CC tumourigenesis. The proteins and their corresponding nucleic acid
 CC sequences can therefore be used for the diagnosis, prevention and
 CC treatment of cervical cancers, particularly cervical intraepithelial
 CC neoplasia (CIN) and squamous intraepithelial lesions (SIL). The sequences
 CC may be used as markers in diagnostic assays to detect cancerous
 CC conditions (e.g. by polymerase chain reaction (PCR) or enzyme linked
 CC immunosorbent assay (ELISA)), to monitor the efficacy of anticancer
 CC therapies and to identify anticancer or carcinogenic compounds. The level
 CC of expression of a marker in a patient sample is compared with the normal
 CC level of expression of the marker in a control non-cervical cancer
 CC sample, whereby a significant difference indicates that the patient is
 CC afflicted with cervical cancer or a pre-malignant condition. The
 CC expression of these proteins may be inhibited by antisense inhibition for
 CC the treatment of cancers. They may be used in this way for the treatment
 CC of cervical intraepithelial neoplasia or squamous intraepithelial
 CC lesions.

Query Match	4.4%;	Score 53;	DB 22;	Length 198;
Best Local Similarity	54.9%;	Pred. No. 0.006;		
Matches	95;	Conservative	0;	Mismatches 78;
				Indels 0;
				Gaps 0;

QY 1038 ggattcttccttcagtttagtccgtatatttatgatattcctacccatttgattgtg 1097
+ | | | | | | | | | | | | | | | | | |
Db 24 gtactcttt 83

QY 1098 tatgatttttcccttaaaatrtataatttcctcaattcgtgaagaaattgaatgat 1157
+ + | ||||| + | |||| + |||| + |||| + |||| +
Db 84 tttttttttttttttttttttttttttttttttttaaaaaaaaanaaa 143

```

QY 1158 attgtacttctgtcaataatagacaagacattcgcaaaaaaaaaa 1210
      | | | | | | | | | | | | | | | | | | | | | |
Db 144 aaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaaa 196

```

RESULT 10

AL11578
ID AL11578 standard; cDNA; 337 BP.

AC AAL11578;

DT 07-DEC-2001 (first entry)

Human breast cancer expressed polynucleotide 4035.

KW Human; breast cancer; cell marker; cytostatic; ss

OS Homo sapiens

PN WO200151628-A2.

PD 19-JUL-2001.

XX

PF 10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077.

PR 24-MAR-2000; 2000US-0192099.

PR 15-MAY-2000; 2000US-0205230.

PR 25-JUL-2000; 2000US-0220534.

PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.

PI Lillie J, Xu Y, Wang Y, Steinmann K;

DR WPI; 2001-451856/48.

PT New peptide useful as a marker for the diagnosis of breast cancer

PS Claim 1; Page 737; 3695pp; English.

CC The invention relates to human breast cancer expressed polynucleotides
CC (AA07544-AA126789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.

Sequence 337 BP; 118 A; 16 C; 6 G; 148 T; 49 other;

Query Match	4.48;	Score 52.8;	DB 22;	Length 337;
Best Local Similarity	54.98;	Pred. No. 0.0076;		
Matches 84;	Conservative 0;	Mismatches 69;	Indels 0;	Gaps 0;

Oy	1058	gtccctgataatttgatatacctaatttggatgtagatgatatttttccttaaa	1117
Db	9	gtactctttttttttttttttaattcttttttttttttttttttttttttttt	68

Qy	1118	tttataattttc	octaatctct	gtaagtaatt	gtaatgata	tttgtacttct	gtcaata	1177
Db	69	tttttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	tttttttttt	128

Qy	1178	atagaacaagacatctgcacaaaaaaaaaaaaa	1210
Db	129	anaaaangnaaaanntnaaaaa	161

RESULT 11

ID ALL10020 standard; CDNA; 394 BP.

AC AAL10020;

DT 07-DEC-2001 (first entry)

DE Human breast cancer expressed polynucleotide 2477.

KW Human; breast cancer; cell marker; cytostatic; ss.

OS Homo sapiens.

PN W0200151628-A2.

PD 19-JUL-2001.

PF 10-JAN-2001; 2001WO-US00798.

PR 14-JAN-2000; 2000US-0176077

PR 24-MAR-2000; 2000US-0192099.

PR

```

PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
XX
PT New peptide useful as a marker for the diagnosis of breast cancer -
PS Claim 1; Page 473; 3695pp; English.
XX
CC The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides are encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
CC
SQ Sequence 394 BP; 218 A; 11 C; 16 G; 93 T; 56 other;
XX
Query Match 4.3%; Score 52.6; DB 22; Length 394;
Best Local Similarity 49.7%; Pred. No. 0.0088;
Matches 91; Conservative 0; Mismatches 92; Indels 0; Gaps 0;
OY 1028 ttaattactagagatttcttcctccagttagtcctgtatttatgatalcttaccta 1087
DB 222 TTTTNTTTTTTTTANNNTTTTATTTATTTTNTTTTTTTTTTTTTTTTTTTT 163
OY 1088 ttgatgtgtatgattttctcctaaattataatttcctaatcttgtaagtaa 1147
DB 162 TAATTTTTTTTTTTANNTTTTNNAAATTTTTTTNNTTTTTTTTTTTTNNAN 103
OY 1148 ttgaatgatatgttaccttcgtcaataatagaacaagacattcgcaaaaaaaaa 1207
DB 102 TTNTTTTTTTNTNTTCCTTAATTTAANNAANAANTNANTNAANNAANAAAAAAA 43
OY 1208 aaa 1210
DB 42 AAA 40
RESULT 12
ID AAS60450 standard; cDNA; 425 BP.
XX
AC AAS60450;
XX
DE 29-JAN-2002 (first entry)
XX
DE Human cancer agent-sensitive marker #181.
XX
KW Human; cancer cell marker; TAXOL; cytosstatic; tumour; carcinoma;
KW squamous cell carcinoma; sarcoma; fibrosarcoma; leukaemia;
KW lymphocytic leukaemia; lymphoma; plasmocytoma; reticulum cell sarcoma;
KW Hodgkin's disease; glioma; ss.
XX
OS Homo sapiens.
XX
PN WO200179556-A2.
XX
PD 25-OCT-2001.
XX
PF 13-APR-2001; 2001WO-US12132.
XX
PR 14-APR-2000; 2000US-197538P.

```

PX	(MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
PA	Lillie J, Brown JL, Bolt A, Van Hufel C;
PI	WPI; 2001-602933/68.
PS	Novel nucleic acid, used as a marker to determine the effectiveness of using TAXOL to treat cancer cell growth in individuals .
XX	Claim 1; Page 221; 527pp; English.
CC	The invention relates to 1046 novel nucleic acids which are used as
CC	markers for determining the sensitivity of a cancer cell to the
CC	anticancer agent TAXOL. Cancer cells can be treated with TAXOL when
CC	they are shown to express one of the 242 sensitivity markers or the
CC	cells are shown not to express one of the 804 resistance markers.
CC	The methods can be used to determine the effectiveness of TAXOL
CC	in the treatment of cancer cell growth in an individual. The markers
CC	can be used as targets in developing anti-cancer agents such as
CC	chemotherapeutic compounds. The markers can also be used as targets in
CC	developing treatments for cancer, particularly those cancers which
CC	display resistance to agents and exhibit expression of the markers. The
CC	anticancer agents developed by the novel method can be used to treat
CC	cancer. Probes based on the markers can be used to detect transcripts or
CC	genomic sequences corresponding to the markers, in the identification of
CC	cells or tissues which mis-express the protein. Cancers which may
CC	be targeted include carcinoma (e.g. squamous cell carcinoma),
CC	sarcoma (e.g. fibrosarcoma) leukaemia (e.g. lymphocytic leukaemia),
CC	lymphoma, plasmacytoma, reticulum cell sarcoma, Hodgkin's disease and
CC	tumours (e.g. glioma). The present sequence is one of the 1046
CC	novel cancer cell markers.
SQ	Sequence 425 BP; 159 A; 41 C; 28 G; 194 T; 3 other;
XX	
Query Match	4.3%; Score 52.4; DB 22; Length 425;
Best Local Similarity	55.4%; Pred. NO. 0.0099;
Matches 98; Conservative	0; Mismatches 79; Indels 0; Gaps 0;
OY	1034 actagattttcttcacggttagcctgtatatttatgatattctactattgat 1093
Dd	1 attt 60
OY	1094 tgtgatgatattttttcctaataaatattatatattcccaatttgttaagtgaattgaat 1153
Dd	61 ttnnaaaaaaaa 120
OY	1154 ggataattgtacctttctgtcataataagaacaagacattcgcaaaaaaaaaaa 1210
Dd	121 aaaaaattttaaaaaaaaaaaaaaaaaaattcccccaaaaaaaaaaaaaa 177
RESULT	13
AAL11441	
ID	AAL11441 standard; cDNA; 426 BP.
AC	AAL11441;
DT	07-DEC-2001 (first entry)
DE	Human breast cancer expressed polynucleotide 3898.
KW	Human; breast cancer; cell marker; cytostatic; ss.
OS	Homo sapiens.
PN	WO200151628-A2.
PD	19-JUL-2001.
PF	10-JAN-2001; 2001WO-US00798.
PX	

PR 14-JAN-2000; 2000US-0176077.
PR 14-MAR-2000; 2000US-0189167.
PR 24-MAR-2000; 2000US-0192099.
PR 29-MAR-2000; 2000US-0193480.
PR 15-MAY-2000; 2000US-0205230.
PR 09-JUN-2000; 2000US-0211315.
PR 25-JUL-2000; 2000US-0220534.
XX
XX
PA (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI Lillie J, Xu Y, Wang Y, Steinmann K;
XX
DR WPI; 2001-451856/48.
PT
XX
XX New peptide useful as a marker for the diagnosis of breast cancer
PS Claim 1; Page 712; 3695pp; English.
XX
XX The invention relates to human breast cancer expressed polynucleotides
CC (AAL07544-AAL26789) and methods of assessing whether a patient is
CC afflicted with breast cancer by examining the correlation between the
CC expression of certain markers and the cancerous state of breast cells.
CC The polynucleotides and encoded polypeptides are potential markers for
CC detecting, diagnosing, monitoring, characterising treating and
CC potentially preventing breast cancer. The polynucleotides and encoded
CC polypeptides are also useful for isolating compounds with cytostatic
CC activity.
XX
XX Sequence 426 BP; 81 A; 57 C; 36 G; 159 T; 93 other;
SQ

```

Query Match          4.3%; Score 52.4; DB 22; Length 426;
Best Local Similarity 50.6%; Pred. No. 0.01;
Matches      86; Conservative    0; Mismatches   84; Indels     0; Gaps     0;

QY  1041 tttcttcacagtttagtcctgtaatttattgtatatcctaactatgtgatgtgtat 1100
      ||||| ||| ||| | ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   45  ttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 104

QY  1101 gattttttccctaaaattttaataatttcctaattccttgtaagtgaattgaatgatatt 1160
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
DB   105 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 164

QY  1161 tgtaccttcgtcaataataagaacacatcgcaaaaaaaaaaaaaaa 1210
      || || || || || || || || || || || || || || || || || || || || ||
DB   165 ttttttnggggnaaaaaaaaannanaaannttttttttttttttttttttttttttttt 214

RESULT  14
ABL32513 standard; DNA; 8805 BP.
XX ABL32513;
AC ABL32513;
XX 
DT 26-MAR-2002 (first entry)
XX 
DE Human immune system associated gene SEQ ID NO: 486.
XX 
KW Human; immune system disease; cytosine methylation; antiasthmatic;
KW antiarteriosclerotic; antianaemic; cytostatic; nootropic;
KW neuroprotective; anti-HIV; anticonvulsant; ophthalmological;
KW antirheumatic; antiarthritic; antidiabetic; antipsoriatic;
KW antiinflammatory; cancer; eye disease; arteriosclerosis; anaemia;
KW acute myeloid leukaemia; Alzheimer's disease; AIDS; epilepsy;
KW neurofibromatosis; rheumatoid arthritis; psoriasis; bowel disease;
KW gene; ds.
XX 
XX Homo sapiens.
OS 
XX 
PN MO200200928-A2.
XX 
PD 03-JAN-2002.
XX 

```

PF 02-JUL-2001; 2001WO-EP07537.
XX
XX 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
PA (EPIG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-130909/17.
DR
XX
XX Nucleic acid comprising fragment of chemically modified gene, useful
PT for diagnosis and treatment of diseases associated with abnormal
PT cytosine methylation -
XX
XX
PS Claim 1; SEQ ID NO 486; 32pp + Sequence Listing; German.
XX
XX The present invention provides a number of human immune system associated
CC genes which are modified by the methylation of cytosines. The sequences
CC can be used in the diagnosis and treatment of immune system disorders,
CC including eye diseases such as retinopathy, neovascular glaucoma and
CC macular degeneration, arteriosclerosis, anaemia, cancer, acute myeloid
CC leukaemia, Alzheimer's disease, AIDS, epilepsy, neurofibromatosis,
CC rheumatoid arthritis, psoriasis and inflammatory/ulcerative bowel
CC diseases. The present sequence is a gene of the invention.
XX
SQ Sequence 8805 BP; 2030 A; 477 C; 2377 G; 3921 T; 0 other;
XX

Query Match	4.3%;	Score 52.4;	DB 24;	Length 8805;
Best Local Similarity	50.4%;	Pred. No. 0.021;		
Matches 128;	Conservative 0;	Mismatches 126;	Indels 0;	Gaps 0
QY 939	tgtaaagatttttgcgtagtcgaagctattttaacagaagtttgcgtattttttccaagaaa	998		
Db 6722	tgttatttaaatttaaataataaataaataaataattttttttgttgagaattt	6781		
QY 999	tctaggacaaggtcaaccctgtgctggcgaaattaactaggaattttcttccaagttag	1058		
Db 6782	ttaagtaaaagtttttttttttgagtggtgtgttttttggaattttatttggtattt	6841		
QY 1059	tcctgtattttatttgatattccttaacctatttgattgtgtatgattttttccctaaat	1118		
Db 6842	ttgtgtattata	6901		
QY 1119	ttataatttcctaattctgtgaagtaattgagatatttgactttctgtcaataa	1178		
Db 6902	gtttatttttttataatcgtttttttttttataattttattttttatatttttttt	6961		
QY 1179	tagaacaagacatt	1192		
Db 6962	ttatttaagtgtt	6975		
RESULT 15				
AAS61101				
ID	AAS61101 standard; DNA: 12592 BP.			
XX				
AC	AAS61101;			
XX				
DT	29-JAN-2002 (first entry)			
XX				
DE	Human gene regulation-associated gene oligonucleotide #56.			
XX				
KW	Human; gene regulation-associated gene; severe combined immunodeficiency;			
KW	cardiac damage; inflammatory response; Haemophilia; Werner syndrome;			
KW	asthma; HDR syndrome; congenital heart defect; Saethre-Chotzen syndrome;			
KW	renal disease; Preeclampsia; cardiac allograft vascular disease;			
KW	colorectal cancer; thyroid cancer; oesophageal cancer; ds; tumour;			
KW	immunostimulant; cardant; antiinflammatory; coagulant; antiasthmatic;			
KW	nephrotropic; gynecological; anti-tumour; immunosuppressive; cyostatic.			
XX				
OS	Homo sapiens.			

Search completed: June 4, 2002, 21:31:08
Job time: 3816 sec

XX WO200177375-A2.
PN 18-OCT-2001.
XX
PD
XX
XX 06-APR-2001; 2001WO-EP03968.
PF
XX
PR 06-APR-2000; 2000DE-1019058.
PR 07-APR-2000; 2000DE-1019173.
PR 30-JUN-2000; 2000DE-1032529.
PR 01-SEP-2000; 2000DE-1043826.
XX
XX (EPiG-) EPIGENOMICS AG.
XX
PI Olek A, Piepenbrock C, Berlin K;
XX
XX WPI; 2002-017470/02.
DR
XX
PT New nucleic acid sequences from chemically modified genes associated
PT with gene regulation, useful for analysing cytosine methylations for
PT diagnosis and therapy of diseases e.g. severe combined immunodeficiency
PT disease
XX
XX
PS Claim 1; SEQ ID No 57; 26pp; English.
XX
CC The invention relates to 224 nucleic acid sequences comprising at least
CC 18 bases of a chemically pretreated gene associated with gene regulation
CC selected from 43 known genes (or complementary sequences). The
CC chemical pretreatment converts cytosine bases unmethylated at the
CC 5-position to uracil or another base with hybridisation behaviour
CC dissimilar to cytosine, to enable analysis of cytosine methylations.
CC The DNA sequences, oligomers (or sets/arrays) and method are
CC useful in the diagnosis of diseases (or predisposition to diseases)
CC associated with gene regulation and in therapy of such diseases, by
CC enabling analysis of the cytosine methylation patterns of such genes,
CC kits are provided. They are especially useful in diagnosis
CC and therapy of e.g. severe combined immunodeficiency disease, cardiac
CC disorders, haemophilia, solid tumours and cancer, Werner syndrome,
CC asthma, HDR syndrome, Saethre-Chotzen syndrome, renal disease,
CC preclampsia, graft versus-host disease. The present sequence is a
CC sequence included in the sequence data for this specification and is
CC associated with the human gene regulation-associated genes.
CC Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic
CC format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 12592 BP; 3783 A; 195 C; 2736 G; 5878 T; 0 other;

Query Match 4.3%; Score 52.2; DB 24; Length 12592;
Best Local Similarity 53.9%; Pred. No. 0.026;
Matches 130; Conservative 0; Mismatches 108; Indels 3; Gaps 1;

QY 939 tgttaagatttttgctagctcaagctattttaacagaagtttggtattttttcagaataa 998
DB 6055 tatttgaagttataaagtaaatgtatttttaagaattttttatatgtttttaaaag 6114
QY 999 tctaggacaaggtcaacctgtgctgagatttaattactagatttttctccagttag 1058
DB 6115 tttagaggtat--ttaaagtaataatattttaataatatattattattattat 6171
QY 1059 tcctgtattttatttgataattcctaactattgattgtgatgatttttccctaaat 1118
DB 6172 tttttaaatagtagtaataattatttttttagttgttttttttttagttgta 6231
QY 1119 ttataaatttccctaattcttgtaagtaattgataattgtactttctgtcaataa 1178
DB 6232 tataagtagtagttatgtgttgaattaaataaacgtaagattatactttaaaattta 6291
QY 1179 t 1179
DB 6292 t 6292

GenCore version 4.5
Copyright (c) 1993 - 2000 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2002, 20:26:47 ; Search time 49.17 Seconds
(without alignments)
6044.672 Million cell updates/sec

Title: US-09-805-919-3
Perfect score: 1210
Sequence: 1 tatattcaattgaaacatg.....ttcgcaaaaaaaaaaaaaa 1210

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 383533 seqs, 122816752 residues

Total number of hits satisfying chosen parameters: 767066

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/2/ina/5A-COMB.seq:*
2: /cgn2_6/ptodata/2/ina/5B-COMB.seq:*
3: /cgn2_6/ptodata/2/ina/6A-COMB.seq:*
4: /cgn2_6/ptodata/2/ina/6B-COMB.seq:*
5: /cgn2_6/ptodata/2/ina/PCTUS-COMB.seq:*
6: /cgn2_6/ptodata/2/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	49.4	4.1	9636	1 US-08-323-170B-1	Sequence 1, Appli
2	49.4	4.1	9636	4 US-08-954-441-1	Sequence 1, Appli
3	45.8	3.8	6243	2 US-09-056-075-1	Sequence 1, Appli
4	44.6	3.7	152331	3 US-09-128-155-16	Sequence 16, Appli
5	44.4	3.7	6243	2 US-09-056-075-1	Sequence 13, Appli
6	44.4	3.7	19124	2 US-08-487-826B-13	Sequence 1, Appli
7	44.2	3.7	3926	2 US-08-731-722-1	Sequence 2, Appli
8	44.2	3.7	3926	2 US-08-731-722-2	Sequence 3, Appli
9	44.2	3.7	3933	2 US-08-731-722-3	Sequence 14, Appli
10	43.8	3.6	2447	2 US-09-014-969-14	Sequence 4, Appli
11	43.6	3.6	1066	1 US-08-157-101A-4	Sequence 1, Appli
12	43.6	3.6	1882	4 US-09-370-253-1	Sequence 33, Appli
13	43.2	3.6	1117	4 US-09-247-373B-33	Sequence 17, Appli
14	43	3.6	176373	3 US-09-128-155-17	Sequence 2, Appli
15	42.8	3.5	5852	1 US-07-867-106-2	Sequence 2, Appli
16	42.6	3.5	2852	3 US-09-027-137-2	Sequence 1, Appli
17	42.4	3.5	2674	4 US-09-817-180-1	Sequence 6, Appli
18	42	3.5	240	1 US-08-628-417-6	Sequence 19, Appli
19	42	3.5	1454	4 US-09-372-422A-19	Sequence 595, App
20	41.8	3.5	658	4 US-08-998-416-595	Sequence 9, Appli
21	41.6	3.4	1172	1 US-07-945-288-9	Sequence 9, Appli
22	41.6	3.4	1172	1 US-08-462-831-9	Sequence 9, Appli
23	41.6	3.4	1172	1 US-08-461-809-9	Sequence 9, Appli
24	41.6	3.4	1172	1 US-08-461-441-9	Sequence 9, Appli
25	41.6	3.4	1172	5 PCT-US93-08518-9	Sequence 9, Appli
26	41.6	3.4	1474	4 US-08-821-994-64	Sequence 64, Appli
27	41.6	3.4	1700	2 US-08-897-340-4	Sequence 4, Appli

C 28	41.6	3.4	1700	3 US-09-252-329-4	Sequence 4, Appli
C 29	41.4	3.4	1582	3 US-08-545-196B-10	Sequence 10, Appli
C 30	41.4	3.4	1582	3 US-08-545-196B-12	Sequence 12, Appli
C 31	40.8	3.4	1134	3 US-09-248-335-29	Sequence 29, Appli
C 32	40.8	3.4	1411	4 US-08-964-127-5	Sequence 5, Appli
C 33	40.8	3.4	1411	4 US-09-496-692-5	Sequence 2, Appli
C 34	40.8	3.4	5852	1 US-07-867-106-2	Sequence 5, Appli
C 35	40.6	3.4	140	1 US-08-628-417-5	Sequence 24, Appli
C 36	40.4	3.3	1493	1 US-08-340-820-24	Sequence 24, Appli
C 37	40.4	3.3	1493	1 US-08-593-535-24	Sequence 4, Appli
C 38	40.4	3.3	3138	1 US-07-867-106-4	Sequence 30, Appli
C 39	40.2	3.3	270	2 US-08-520-678A-30	Sequence 30, Appli
C 40	40.2	3.3	270	4 US-08-897-126-30	Sequence 1, Appli
C 41	40.2	3.3	3763	1 US-07-792-865D-1	Sequence 10, Appli
C 42	40	3.3	1349	2 US-08-676-782-10	Sequence 1, Appli
C 43	40	3.3	8920	2 US-08-446-855A-1	Sequence 1, Appli
C 44	40	3.3	8920	4 US-09-150-741-1	Sequence 36, Appli
C 45	39.6	3.3	665	2 US-08-883-795A-36	

ALIGNMENTS

RESULT 1
US-08-323-170B-1
Sequence 1, Application US/08323170B
Patent No. 5733772
GENERAL INFORMATION:
APPLICANT: Williamson, Kim C.
TITLE OF INVENTION: Cloning and Expression of Plasmodium
TITLE OF INVENTION: falciparum Transmission-Blocking Target Antigen, Pf5230
NUMBER OF SEQUENCES: 4
CORRESPONDENCE ADDRESSES:
ADDRESS: Townsend and Townsend and Crew LLP
STREET: Two Embarcadero Center, 8th Floor
CITY: San Francisco
STATE: California
COUNTRY: USA
ZIP: 94111-3834
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/323,170B
FILING DATE: 13-OCT-1994
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/010,409
FILING DATE: 29-JAN-1993
ATTORNEY/AGENT INFORMATION:
NAME: Quine, Jonathan A.
REGISTRATION NUMBER: P-41,261
REFERENCE/DOCKET NUMBER: 015280-113100US
TELECOMMUNICATION INFORMATION:
TELEPHONE: (415) 576-0200
TELEFAX: (415) 576-0300
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 9636 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
FEATURE:
NAME/KEY: CDS
LOCATION: 149..9556
US-08-323-170B-1
Query Match 4.1%; Score 49.4; DB 1; Length 9636;

QY	1080	cttaacctatttgatctgctgatgatcttcttcccttaaaaaatttatacaatttcctaattctt	1139
Db	3318	TAAATATGTTTATTCAAATATTTATCTTTTTCFAAATTATATATATATTTTATTATATTTA	3377
QY	1140	gtaaagtaattgtaatggaataattgtactcttcgcacaataatagaacaagacatttcgc meta	1199
Db	3378	TTATTATATATATTTTATTTTAACTTCTTTCTAACAAGCTATTAAAGAAGAACTTAATA	3437
QY	1200	aaaaaaaaaa	1208
Db	3438	ATAAAAAACA	3446

```

RESULT      4
US-09-128-155-16/C
; Sequence 16, Application US/09128155
; Patent No. 6117654
;
GENERAL INFORMATION:
; APPLICANT: Pan, Yang
;
TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
;
FILE REFERENCE: 09404/052001
;
CURRENT APPLICATION NUMBER: US/09/128,155
;
CURRENT FILING DATE: 1998-08-03
;
EARLIER APPLICATION NUMBER: US 60/091,650
;
EARLIER FILING DATE: 1998-07-02
;
EARLIER APPLICATION NUMBER: US 60/054,646
;
EARLIER FILING DATE: 1997-08-04
;
NUMBER OF SEQ ID NOS: 18
;
SOFTWARE: FastSeq for Windows Version 3.0
;
SEQ ID NO 16
;
LENGTH: 152331
;
TYPE: DNA
;
ORGANISM: Homo sapiens
;
FEATURE:
;
NAME/KEY: misc_feature
;
LOCATION: (1)...(152331)
;
OTHER INFORMATION: n = A,T,C or G
US-09-128-155-16

```

	Query Match	3.7%	Score 44.6;	DB 3;	Length 152331;
	Best Local Similarity	48.6%;	Pred. No. 0.069;		
	Matches 122;	Conservative	0;	Mismatches 129;	Indels 0;
				Gaps	0;
QY	903	agacggaattgcatacatccctccgatgaacagacigttaagatttttgctgaagc	962		
Db	23610	AGATAGCGCTTAACCAATAGAACACATCTGTATATAAATGTAGATTAGAGTTAACTTTC	23551		
QY	963	tattttaacagaagttgtgtatrttttttcagaaatctagacaaggtcaacctgtgct	1022		
Db	23550	CAATCACACTAATTCGTTTTATGTGAAALAAAGGAATGAACCTGTTCCATCGCTGGTGGAAAGAT	23491		
QY	1023	ggcgattaatctactagagattttcttcttcacggttagtccgtlatltaattgataattctt	1082		
Db	23490	AGAGATTAATTTTAGAGGTTGTGCGTGTGTTTGGGATCTGTTTCTTTAAAAATTGT	23431		
QY	1083	acctatttgatttgtaataatttttcccttaaaatttataaatttccctaattcttgta	1142		
Db	23430	AAATATGTACTGTGTGAATGATTTTTTAAATGATTTTACCATTTTTGGAAAGGTAATTT	23371		
QY	1143	agtaattgaat	1153		
Db	23370	AATGATAGAAAT	23360		

RESULT 5
US-09-056-075-1/c
; Sequence 1, Application US/09056075
; Patent No. 595368
; GENERAL INFORMATION:
; APPLICANT: Johnson, Eric A.
; APPLICANT: Bradshaw, Marite

```

1  APPLICANT:  Rood, Julian
2  TITLE OF INVENTION:  Expression System for Clostridium
3  TITLE OF INVENTION:  Species
4  NUMBER OF SEQUENCES:  2
5  CORRESPONDENCE ADDRESS:
6  ADDRESSEE:  Quarles & Brady
7  STREET:  1 South Pinckney Street
8  CITY:  Madison
9  STATE:  WI
10 COUNTRY:  US
11 ZIP:  53701-2113
12
13 COMPUTER READABLE FORM:
14 MEDIUM TYPE:  Floppy disk
15 COMPUTER:  IBM PC compatible
16 OPERATING SYSTEM:  PC-DOS/MS-DOS
17 SOFTWARE:  PatentIn Release #1.0, Version #1.30
18 CURRENT APPLICATION DATA:
19 APPLICATION NUMBER:  US/09/056,075
20 FILING DATE:
21 CLASSIFICATION:
22 ATTORNEY/AGENT INFORMATION:
23 NAME:  Seay, Nicholas J.
24 REGISTRATION NUMBER:  27386
25 REFERENCE/DOCKET NUMBER:  960296.95238
26 TELECOMMUNICATION INFORMATION:
27 TELEPHONE:  608-251-5000
28 TELEFAX:  608-251-9166
29 INFORMATION FOR SEQ ID NO:  1:
30 SEQUENCE CHARACTERISTICS:
31 LENGTH:  6243 base pairs
32 TYPE:  nucleic acid
33 STRANDEDNESS:  double
34 TOPOLOGY:  linear
35 MOLECULE TYPE:  DNA (genomic)
36 FEATURE:
37 NAME/KEY:  misc_feature
38 LOCATION:  3770..4013
39 OTHER INFORMATION:  /note= "RP4 origin of DNA transfer (oriT) from
40 US-09-056-075-1

```

Query Match	3.78;	Score 44.4;	DB 2;	Length 6243;
Best Local Similarity	51.5%;	Pred. No. 0.022;		
Matches 102;	Conservative	0;	Mismatches 96;	Indels 0;
				Gaps 0;
QY 975 agttgtgataatttttcagaaatcctagacaagctcaacctgtgctgcgatnaatta				
Db 1475 ATTTTATATATCTTTCTTCTCAAGATATATATATAAATAATTTTTCACACTTAA				
QY 1035 ctagattcttccttcocagtttagtcctgtattttattgataattcctacatttgatt				
Db 1415 AATAAATAATATTTTATATTTTATATTTTATTTTATATTTTATTTTATTTTAT				
QY 1095 gtgtatgattttttccctaaatttataaatttccctaattcttgaagtaattgaaatg				
Db 1355 TTTTATATTTTATATTTTATATATTTTATTTTATTTTATATTTTATTTTAT				
QY 1155 gataattgtactttctgt 1172				
Db 1295 TTATTTTATATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTAT				

```

RESULT      6
US-08-487-826B-13/c
; Sequence 13, Application US/08487826B
; Patent No. 5993827
;
; GENERAL INFORMATION:
;
; APPLICANT:  Sim, Kim L.
;
; APPLICANT:  Chitnis, Chetan
;
; APPLICANT:  Miller, Louis H.
;
; APPLICANT:  Peterson, David S.
;
; APPLICANT:  Su, Xin-zhaun

```



```

; APPLICANT: Wellens, Thomas E.
; TITLE OF INVENTION: BINDING DOMAINS FROM PLASMODIUM VIVAX
; TITLE OF INVENTION: AND PLASMODIUM FALCIPARUM ERYTHROCYTE BINDING PROTEINS
; NUMBER OF SEQUENCES: 45
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe Martens Olson & Bear
; STREET: 620 Newport Center Drive 16th Floor
; CITY: Newport Beach
; STATE: California
; COUNTRY: US
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION NUMBER: US/08/487,826B
; APPLICATION NUMBER: US/08/487,826B
; FILING DATE: 10-SEP-1993
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelson, Ned
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: NIH121.001CP1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 235-8550
; TELEFAX: (619) 235-0176
; INFORMATION FOR SEQ ID NO: 13:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 19124 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: CDNA
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; US-08-487-826B-13

Query Match          3.7%; Score 44.4; DB 2; Length 19124;
Best Local Similarity 48.4%; Pred. No. 0.034;
Matches 123; Conservative 0; Mismatches 131; Indels 0; Gaps 0;

QY 929 tgaacagaactgttaagatttttgcctagctcaagctattttaacagaagttgtgtattt 988
DB 15977 TGCATGTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15918

QY 989 ttccagaaatcagagcaagtcacactgtgctgcgaattaaactagagattttctt 1048
DB 15917 TTTCAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 15858

QY 1049 tccagtttagctctgtattttatctatcttcttaccatttgattgtgtatgattttt 1108
DB 15857 TTTATTTAATTAATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTT 15798

QY 1109 tcctlaaaatttataattttcctaattcttgtaagtaattgaatgatattgtacttt 1168
DB 15797 TTATGTATATATTTTATTTTAAACAATTTTAAATTTTATTTTATTTTATTTATGATATAT 15738

QY 1169 ctgtcaataataga 1182
DB 15737 TTTATTTTAAATATA 15724

RESULT 7
US-08-731-722-1
; Sequence 1, Application US/08731722
; Patent No. 5961971
; GENERAL INFORMATION:
; APPLICANT: Martin, Frank N.
; TITLE OF INVENTION: Biocontrol of Fungal Soilborne Pathogens
; TITLE OF INVENTION: by Pythium oligandrum
; NUMBER OF SEQUENCES: 8
```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION NUMBER: US/08/731,722
; APPLICATION NUMBER: US/08/731,722
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF-161
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 1982-24
; US-08-731-722-1

Query Match          3.7%; Score 44.2; DB 2; Length 3926;
Best Local Similarity 51.2%; Pred. No. 0.021;
Matches 103; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY 974 aagttgtgtattttttcagaaatcagagcaagtcacactgtgctgcgataatt 1033
DB 1251 AATGTGAGTGATAGTAGCAAGATCGTTTATTTTGAACCTATCGAGTTTATTTT 1310

QY 1034 actagattttcttccagtttagctctgtattttatctatcttaccatttgat 1093
DB 1311 CTTTGAATTTTCTAATATTTTATCTCTATTTTGTATCGTTTGTCTACTTATATA 1370

QY 1094 tgtgatgatatttttccctaaatttataattttcctaattcttgtaagtaattgat 1153
DB 1371 TTAATAAAATTTTTCCTAATATTTTCAAAATTTTGATTATTTTGTATATATTTTCTT 1430

QY 1154 ggaatttgacttctgtca 1174
DB 1431 AAATTTCTAAGCTTTAGCA 1451

RESULT 8
US-08-731-722-2
; Sequence 2, Application US/08731722
; Patent No. 5961971
; GENERAL INFORMATION:
; APPLICANT: Martin, Frank N.
; TITLE OF INVENTION: Biocontrol of Fungal Soilborne Pathogens
; TITLE OF INVENTION: by Pythium oligandrum
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
```

```
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,722
; FILING DATE:
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF-161
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3926 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; STRAIN: 1985-5
; US-08-731-722-2

Query Match          3.7%; Score 44.2; DB 2; Length 3926;
Best Local Similarity 51.2%; Pred. No. 0.021;
Matches 103; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY  974 aagtttgatgattttttcagaaatctagagcaaggtcaacctgtgctggcgattaatt 1033
    |||||
DB  1251 AATGTGAGTGAGTAGTAGCAAGATCGGTTTATTGGAACCTCATCGAGTTATT 1310

QY  1034 actagagattttcttccagtttagtcctgtatttattgataattcttaacctattgat 1093
    |||||
DB  1311 CTTGAATTTTCTAATAATTATTCCTCTATTTGTTATCGTTTGTCTACTTATATA 1370

QY  1094 tgtgatgattttttccctaaatttataattttccctaattcttgtaagtaattgaat 1153
    |||||
DB  1371 TTAATAAATTTTTCCTAATATTTCACAAATTTGATTATTGTTGATATATTTCCT 1430

QY  1154 ggataattgtacttctgtca 1174
    |||||
DB  1431 AAATTTCTACTCTTTAGCA 1451

RESULT 9
US-08-731-722-3
; Sequence 3, Application US/08731722
; Patent No. 5961971
; GENERAL INFORMATION:
; APPLICANT: Martin, Frank N.
; TITLE OF INVENTION: Biocontrol of Fungal Soilborne Pathogens
; TITLE OF INVENTION: by Pythium oligandrum
; NUMBER OF SEQUENCES: 8
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Saliwanchik & Saliwanchik
; STREET: 2421 N.W. 41st Street, Suite A-1
; CITY: Gainesville
; STATE: FL
; COUNTRY: US
; ZIP: 32606-6669
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/731,722
; FILING DATE:
; CLASSIFICATION: 424
```

```
; ATTORNEY/AGENT INFORMATION:
; NAME: Whitlock, Ted W.
; REGISTRATION NUMBER: 36,965
; REFERENCE/DOCKET NUMBER: UF-161
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 352-375-8100
; TELEFAX: 352-372-5800
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3933 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; INDIVIDUAL ISOLATE: 1986-41
; US-08-731-722-3

Query Match          3.7%; Score 44.2; DB 2; Length 3933;
Best Local Similarity 51.2%; Pred. No. 0.021;
Matches 103; Conservative 0; Mismatches 98; Indels 0; Gaps 0;

QY  974 aagtttgatgattttttcagaaatctagagcaaggtcaacctgtgctggcgattaatt 1033
    |||||
DB  1255 AATGTGAGTGAGTAGTAGCAAGATCGGTTTATTGGAACCTCATCGAGTTATT 1314

QY  1034 actagagattttcttccagtttagtcctgtatttattgataattcttaacctattgat 1093
    |||||
DB  1315 CTTGAATTTTCTAATAATTATTCCTCTATTTGTTATCGTTTGTCTACTTATATA 1374

QY  1094 tgtgatgattttttccctaaatttataattttccctaattcttgtaagtaattgaat 1153
    |||||
DB  1375 TTAATAAATTTTTCCTAATATTTCACAAATTTGATTATTGTTGATATATTTCCT 1434

QY  1154 ggataattgtacttctgtca 1174
    |||||
DB  1435 AAATTTCTACTCTTTAGCA 1455

RESULT 10
US-09-014-969-14/c
; Sequence 14, Application US/09014969
; Patent No. 5965397
; GENERAL INFORMATION:
; APPLICANT: Jacobs, Kenneth
; APPLICANT: McCoy, John M.
; APPLICANT: Lavallee, Edward R.
; APPLICANT: Racie, Lisa A.
; APPLICANT: Merberg, David
; APPLICANT: Treacy, Maurice
; APPLICANT: Spaulding, Vikki
; APPLICANT: Agostino, Michael J.
; TITLE OF INVENTION: SECRETED PROTEINS AND POLYNUCLEOTIDES
; TITLE OF INVENTION: ENCODING THEM
; NUMBER OF SEQUENCES: 32
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Genetics Institute, Inc.
; STREET: 87 Cambridgepark Drive
; CITY: Cambridge
; STATE: MA
; COUNTRY: U.S.A.
; ZIP: 02140
; COMPUTER READABLE FORM:
; MEDIUM TYPE: FLOPPY disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/014,969
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
```

```

; NAME: Sprunger, Suzanne A.
; REGISTRATION NUMBER: 41,323
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (617) 498-8284
; TELEFAX: (617) 876-5851
; INFORMATION FOR SEQ ID NO: 14:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2447 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
;
US-09-014-969-14

```

Query Match	3.68;	Score 43.8;	DB 2;	Length 2447;
Best Local Similarity	52.58;	Pred. No. 0.022;		
Matches 96;	Conservative	0;	Mismatches 87;	Indels 0;
				Gaps 0;

QY	1028	ttaatctactaggattttctcttccagtttagtcctgtatttlatattgtaattcttaccta	1087
Db	2385	TT	2326
QY	1088	tttgatgtgtatgatctttcttcccttaaaaatttataaatttctccaaattcttgytaagtaa	1147
Db	2325	TT	2266
QY	1148	ttgaatggaatatgttactttctgtcgaataatagaacaagacatcgcgaaaaaaanaaa	1207
Db	2265	TTTTTTTTTTTTTTTTTTTTTTCTCCCAAAAGGTTTATTGTGCACATTTAAAGTACAAAT	2206
QY	1208	aaa 1210	
Db	2205	CAA 2203	

RESULT 11
US-08-157-101A-4/c
; Sequence 4, Application US/08157101A
; Patent No. 5808032
; GENERAL INFORMATION:
; APPLICANT: KURIHARA, TATSUYA
; APPLICANT: MATSUKURA, SHIGEKAZU
; APPLICANT: TSURUOKA, NOBUO
; APPLICANT: ARIMA, KENJI
; APPLICANT: NISHIHARA, TATSURO
; TITLE OF INVENTION: ANTI-HBS ANTIBODY GENES AND EXPRESSION
; TITLE OF INVENTION: PLASMIDS THEREFOR
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO
; STREET: 1100 NEW YORK AVENUE, N.W.
; CITY: WASHINGTON
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/157,101A
; FILING DATE: 05-APR-1994
; CLASSIFICATION: 530
; ATTORNEY/AGENT INFORMATION:
; NAME: TITUS, MARLANA K
; REGISTRATION NUMBER: 35843
; REFERENCE/DOCKET NUMBER: 9437/204199
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 202-861-3711
; TELEFAX: 202-822-0944
; TELEX: 6714627 CUCH

```

; INFORMATION FOR SEQ ID NO: 4:
;     SEQUENCE CHARACTERISTICS:
;         LENGTH: 1066 base pairs
;         TYPE: nucleic acid
;         STRANDEDNESS: single
;         TOPOLOGY: linear
;     MOLECULE TYPE: DNA (genomic)
;     US-08-157-101A-4

```

Query Match	3.68;	Score 43.6;	DB 1;	Length 1066;
Best Local Similarity	56.28;	Pred. No. 0.018;		
Matches	82;	Conservative	0;	Mismatches 64;
			Indels	0;
			Gaps	0;

QY 1041 ttttctttccagttttagtcctgtattttattgataattcttaaccatttgattgtgtat 1100
 ||||| ||| ||| | ||||| ||| | ||| ||| ||| |||
 Db 1037 TT 978
 QY 1101 gattttttcccttaaaattataattttcccaattctgtgaagtaattgaaatgatatt 1160
 ||||| || ||||| ||||| ||||| ||||| ||||| |||||
 Db 977 TTGCAGAAGATTCACTTATTTA 918
 QY 1161 tgtactttctctgtcaataataagaacaa 1186
 | | | | | ||| | ||| | ||| | |||
 Db 917 TTCAATTCTCTCCACACATTAGCATTA 892

```

RESULT 12
US-09-370-253-1/c
; Sequence 1, Application US/09370253
; Patent No. 6165792
; GENERAL INFORMATION:
; APPLICANT: Allen, Stephen M.
; APPLICANT: Sakai, Hajime
; APPLICANT: Thorpe, Catherine J.
; TITLE OF INVENTION: Amino Acid Transporters
; FILE REFERENCE: BB-1200
; CURRENT APPLICATION NUMBER: US/09/370,253
; CURRENT FILING DATE: 1999-08-09
; EARLIER APPLICATION NUMBER: 60/097,222
; EARLIER FILING DATE: August 20, 1998
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Microsoft Office 97
; SEQ ID NO 1
; LENGTH: 1882
; TYPE: DNA
; ORGANISM: Hordeum vulgare
US-09-370-253-1

```

Query Match	3.6%;	Score 43.6;	DB 4;	Length 1882;
Best Local Similarity	56.2%;	Pred. No. 0.022;		
Matches	82;	Conservative	0;	Mismatches 64;
			Indels	0;
			Gaps	0

QY	1041	tcttcctccacagtttaagtcctgataatttattgatatactcctaaccatttgatgtgtat	11000
Db	1880	TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT	18222
QY	1101	gattttttctocctaaatttataatttctoctaatcttgtaagtaattggaatgatatatt	11600
Db	1820	TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTTTT TTTTGAAGAAAAGAACCCATTACA	17611
QY	1161	tgtactttctgccaataatagaaccaa	1186
Db	1760	TTGCGGAATTAATCTATGCACAGAACAA	1735

RESULT 13
US-09-247-373B-33/c
; Sequence 33, Application US/09247373B
; Patent No. 6168954
; GENERAL INFORMATION:
; APPLICANT: MCGONIGLE, BRIAN

```

1  APPLICANT: O'KEEFE, DANIEL
2  TITLE OF INVENTION: SOYBEAN GLUTATHIONE-S-TRANSFERASE ENZYMES
3  FILE REFERENCE: CL-1108-A
4  CURRENT APPLICATION NUMBER: US/09/247,373B
5  CURRENT FILING DATE: 1999-02-10
6  PRIOR APPLICATION NUMBER: 08/924,747
7  PRIOR FILING DATE: 1997-09-05
8  NUMBER OF SEQ ID NOS: 56
9  SOFTWARE: Microsoft Office 97
10 SEQ ID NO 33
11 LENGTH: 1117
12
13 TYPE: DNA
14
15 ORGANISM: SOYBEAN
16
17 FEATURE:
18
19 NAME/KEY: unsure
20 LOCATION: (1101)
21 OTHER INFORMATION: M=A OR C
22
23 NAME/KEY: unsure
24 LOCATION: (1104)
25 OTHER INFORMATION: M=A OR C
26
27 NAME/KEY: unsure
28 LOCATION: (1116)
29 OTHER INFORMATION: N=G or A or T or C
30
31 US-09-247-373B-33

```

Query Match	3.68;	Score 43.2;	DB 4;	Length 1117;
Best Local Similarity	57.48;	Pred. No. 0.023;		
Matches	78;	Conservative	0;	Mismatches 58;
			Indels	0;
			Gaps	0;

QY	1028	ttaatcctaggaatttctctccagtttagtcctgatttatttgataattcttaccta	1087
Db	1110	TTTTTAKKTKTTTTTGGTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT	1051
QY	1088	tttgatcgtgatgatattttctccttaaaatttataaatttccctaattcttgaagtaa	1147
Db	1050	TTAAAGAA	991
QY	1148	ttgaatcgataatttgt	1163
Db	990	TTGAATTATGATTTTAT	975

```

RESULT 14
US-09-128-155-17
; Sequence 17, Application US/09128155
; Patent No. 6117654
; GENERAL INFORMATION:
; APPLICANT: Pan, Yang
; TITLE OF INVENTION: NOVEL MOLECULES OF TANGO-77 RELATED PROTEIN FAMILY
; TITLE OF INVENTION: AND USES THEREOF
; FILE REFERENCE: 09404/052001
; CURRENT APPLICATION NUMBER: US/09/128,155
; CURRENT FILING DATE: 1998-08-03
; EARLIER APPLICATION NUMBER: US 60/091,650
; EARLIER FILING DATE: 1998-07-02
; EARLIER APPLICATION NUMBER: US 60/054,646
; EARLIER FILING DATE: 1997-08-04
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 176373
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)..(176373)
; OTHER INFORMATION: n = A,T,C or G
US-09-128-155-17

```

Query Match	3.6%;	Score 43;	DB 3;	Length 176373;
Best Local Similarity	48.2%;	Pred. No. 0.19;		

Matches	121;	Conservative	0;	Mismatches	130;	Indels	0;	Gaps	0;
QY	903	agacggaatttgcatacatcctccgagtgaacagaactgttaagattttgtctagtcacg	962						
Db	20806	agatagcgttaacaataagacaacatctcgtatataaatgtgttaaatlaggttaattcttc	20865						
QY	963	tattttaacagagaatttgtgtatttttttcagaaatctagacaaggtcaacctgtgct	1022						
Db	20866	caatcacataatctcgttttatgtgtgaaaaaaggaaatgtaactgtttccatgtctgtgtgaaagat	20925						
QY	1023	ggcgattaatctactagaatttttctctccagtttagtcctgtattttatttgatatcttc	1082						
Db	20926	agagattatttcttagagggtttgtcgtctgtgttttgagattcgttttctttaaattgtl	20985						
QY	1083	acctattgattgtgtatgatattttttccttaanaattttataatttccctaattctgtta	1142						
Db	20986	aaatatgtactcgtgtgatgtatttttttaanaatgatttttaacatttttggaaagggtattt	21045						
QY	1143	agtaattgaaat	1153						
Db	21046	aatgataagaat	21056						

RESULT 15
US-07-867-106-2
; Sequence 2, Application US/07867106
; Patent No. 5389526

```

GENERAL INFORMATION:
APPLICANT: Slade, Martin B
APPLICANT: Chang, Andy C M
APPLICANT: Williams, Keith L
TITLE OF INVENTION: Improved Plasmid Vectors for Cellular
TITLE OF INVENTION: Slime Moulds of the Genus Dictyostelium
NUMBER OF SEQUENCES: 19
CORRESPONDENCE ADDRESS:
ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5389526ris
STREET: One Liberty Place 46th Floor
CITY: Philadelphia
STATE: PA
COUNTRY: USA
ZIP: 19103
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/07/867,106
FILING DATE: 19920625
PRIOR APPLICATION DATA:
APPLICATION NUMBER: AU PJ 7187
APPLICATION NUMBER: PCT/AU90/00530
FILING DATE: 02-NOV-1989
ATTORNEY/AGENT INFORMATION:
NAME: Feeney, Joanne Longo
REGISTRATION NUMBER: 35,134
REFERENCE/DOCKET NUMBER: RICE-0002
TELECOMMUNICATION INFORMATION:
TELEPHONE: 215-568-3100
TELEFAX: 215-568-3439
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 5852 base pairs
TYPE: NUCLEIC ACID
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: DNA (genomic)
ANTI-SENSE: NO
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038
FEATURE:
NAME/KEY: CDS
LOCATION: 2378..5038

```


GenCore version 4.5
Copyright (c) 1993 - 2000 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: June 4, 2002, 20:24:37 ; Search time 1589.3 Seconds
(Without alignments)
10275.792 Million cell updates/sec

Title: US-09-805-919-3
Perfect score: 1210
Sequence: 1 tatattcaattgaaacatg.....ttcgcacaaaaaaaaaaaaa 1210

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 13736207 seqs, 6748477542 residues
Total number of hits satisfying chosen parameters: 27472414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:
1: em_estda:*
2: em_esthum:*
3: em_estlin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_estl:*
10: gb_est2:*
11: gb_htc:*
12: gb_gss:*
13: em_gss_hum:*
14: em_gss_hlv:*
15: em_gss_pln:*
16: em_gss_vrt:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	349	28.8	639	10	BI434492
2	335.4	27.7	538	9	AI775018
3	301.6	24.9	533	9	AI897599
4	238.4	19.7	432	9	AI490332
5	70	5.8	654	10	BF645984
6	69.8	5.8	498	10	BF633219
7	67.8	5.6	603	9	AM695109
8	67.6	5.6	329	9	AL513719
9	65.8	5.4	1045	12	CNS03YE4
10	65	5.4	625	10	BF637888
11	65	5.4	945	10	BM358135
12	64.8	5.3	945	10	BM358135
13	63.6	5.3	638	9	AL513901
14	63.6	5.3	991	12	CNS0012D
15	62.6	5.2	353	9	AL515235
16	62.6	5.2	503	9	AL513809
17	62.6	5.2	546	10	BF643246

18	62.4	5.2	660	10	BG591192	BG591192	EST499034
19	62.4	5.2	947	12	CNS015IM	AL105496	Drosoph11
20	61.4	5.1	379	9	AL514359	AL514359	AL514359
21	61	5.0	425	9	AL514791	AL514791	AL514791
22	60.8	5.0	633	9	AL513979	AL513979	AL513979
23	60.6	5.0	365	9	AL515373	AL515373	AL515373
24	60.4	5.0	329	9	AL514627	AL514627	AL514627
25	60	5.0	1043	12	CNS0145P	AL103735	Drosoph11
26	59.6	4.9	319	9	AL513977	AL513977	AL513977
27	59.6	4.9	951	12	AZ672893	AZ672893	ENTUT74TF
28	59.2	4.9	344	9	AL513975	AL513975	AL513975
29	59.2	4.9	894	12	BH137168	BH137168	ENTOK14TR
30	59	4.9	523	9	AL514015	AL514015	AL514015
31	58.8	4.9	712	9	BE038709	BE038709	AB04D10 A
32	58.4	4.8	1101	12	CNS003DQ	AL064580	Drosoph11
33	58.2	4.8	309	9	AI490927	AI490927	EST241636
34	58.2	4.8	468	9	AI489859	AI489859	EST248198
35	58.2	4.8	470	9	AI490270	AI490270	EST248596
36	58.2	4.8	886	12	BH177277	BH177277	008_L_22-
37	58.2	4.8	886	12	CNS07JUX	AL614235	T3 end of
38	58.2	4.8	941	10	BM415213	BM415213	OP20285 M
39	58	4.8	344	9	AL513975	AL513975	AL513975
40	58	4.8	363	9	AL514473	AL514473	AL514473
41	58	4.8	956	12	CNS04WDT	AL310250	Tetraodon
42	57.6	4.8	392	9	AL514511	AL514511	AL514511
43	57.6	4.8	449	9	AL513999	AL513999	AL513999
44	57.6	4.8	920	12	CNS0062R	AL061710	Drosoph11
45	57.6	4.8	954	12	BH162327	BH162327	ENTR037TR

ALIGNMENTS

RESULT 1
BI434492
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BI434492 639 bp mRNA linear EST 21-AUG-2001
EST537253 P. infestans-challenged leaf solanum tuberosum CDNA clone
PPCBR09 5' sequence, mRNA sequence.
BI434492
BI434492.1 GI:15259182
EST.
potato.
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
1 (bases 1 to 639)
Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chienlingo,A.,
Bougri,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.
Generation of ESTs from Potato Leaves Challenged with Phytophthora
infestans, Compatible Interaction
Unpublished (2000)
Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdna@resgen.com
Seq primer: M13F-R.
Location/Qualifiers
1. 639
/organism="Solanum tuberosum"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="PPCBR09"
/clone_lib="P. infestans-challenged leaf"
/tissue_type="leaf"
/dev_stage="6 week old"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Whole plants were challenged with 450,000
sporangia/ml P. infestans US-1(US 940501) in Biotron
(Madison, Wisconsin). Leaf tissue was collected at 1, 2,
5, 12, and 24 hours post-challenge and frozen in liquid
nitrogen immediately upon removal. Kennebec plants showed

no signs of HR. Kathadlin plants (susceptible to P. infestans US-1) were used as controls and showed infection. NOTE: We cannot exclude the possibility that this sequence is actually derived from Phytophthora rather than potato."

BASE COUNT 200 a 144 c 101 g 194 t
ORIGIN

Query Match 28.8%; Score 349; DB 10; Length 639;
Best Local Similarity 76.9%; Pred. No. 3.1e-48;
Matches 485; Conservative 0; Mismatches 125; Indels 21; Gaps 4;

QY 1 tatattcaattgaaacatgttgacaagagggctcttctcgcttggtgttcttacttg 60
DB 9 TATATTCAATTGACAAATGTCCATACCAAAATTCATTTTGGCTAGTGTCTTCTCCCTTG 68
QY 61 tgaacatcataagcagttcttaagcgcaagataatt--tctcaatgtgtcccttcttct 117
DB 69 TTTCATTTCATACAGAGCTCTATAGCCCCAAATAGTAGCTCTCATGTGTCTCTTCTCT 128
QY 118 gcggtgatatt--caataaaattccctccgactgtagactgatacccgagcattgtg 174
DB 129 GTGGCGATATTGCGCAATATAAATTCCTCCCTCCGATTGAGGACTGATCTCGAGCATTTG 188
QY 175 gtagacgcgagatagagctcgattgccaagaacaacaaacgcgtgttcaattacaatcca 234
DB 189 GCAACATGATACGAACTCGATTGCCAGAACATCAAACTTCTTACACTTACAGTCCA 248
QY 235 gaatttcgacgtacaggaatttaactacagaagctactcaataagcctacttgatctg 294
DB 249 GAAATTTCTACGTGACGAGAAATTAACACTACAGCTACATGATAGGCTACTCGATCTG 308
QY 295 gcctaaatgatcagagagaanaattgcacagattttccaataacacagggcaagttat--g 351
DB 309 GCTTAAAGATCAGAACGAAATTTGCTGTGTTTCCCTGATTTACAGAGCAAAATTATTAG 368
QY 352 atgccatgactagccaatcttgaatgggttcgtgttlaacaatgatatcaactatgtca 411
DB 369 GAGGCTTAAGTAGTAATTTTCCAAATGATTCATATTAAACAATAGCATCAACTATGTCA 428
QY 412 actgtctagctcctatcaaatlogtcacaglatatloctacaagttttgtagcaaaaatt 471
DB 429 ACTGTGAACTCCTATCAACTCGTCACAATACATTCCAACGACTTTTGTAGAACAAATA 488
QY 472 caa-----cgggttttagctacctgtgcataagaagaataattgcaagcttcg 519
DB 489 CAACCTTCGCCCAACGCCAATTTTAGTTATCTGTCTTAAAGAAATATGGCAAGCTTCGG 548
QY 520 attggctggcggtcgttaagggttgaactgttgcattgtcctctgtccacagcatttcat 579
DB 549 AATTGGAAACCGGTTGCAAGGTTGAAACTGTGGCGTGGTCTTCAGCTCCAGGTATTTTTA 608
QY 580 caacaagtcgtctacgttatacaagcacaca 610
DB 609 CAAACAAGTCGCCCTTTGTTAGGAAGCACTCA 639

RESULT 2
AI775018 538 bp mRNA linear EST 18-MAY-2001
LOCUS AI775018 tomato resistant, Cornell Lycopersicon esculentum cDNA
DEFINITION EST256118 tomato resistant, Cornell Lycopersicon esculentum cDNA
ACCESSION AI775018
VERSION AI775018
KEYWORDS AI775018.1 GI:5273059
SOURCE EST.
ORGANISM tomato.
LYCOPERSICON ESCULENTUM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 538)

AUTHORS D' Ascenzo, M., He, X., Lyman, J., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman,
C.L., Nierman, W., Fraser, C.M., Venter, J.C., Tanksley, S.D.,
Giovannoni, J.J. and Martin, G.B.
TITLE Generation of ESTs from Pseudomonas resistant tomato
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

FEATURES

source

location/Qualifiers
1..538
/organism="Lycopersicon esculentum"
/cultivar="R11-12 (35S):Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLER14G5"
/clone_lib="tomato resistant, Cornell"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLER - Tomato Pseudomonas Resistant EST Library.
Directionally cloned cDNAs inserted into pBluescript SK(-)
at 5' end with EcoRI and 3' end with XhoI site."

BASE COUNT 145 a 126 c 107 g 160 t
ORIGIN

Query Match 27.7%; Score 335.4; DB 9; Length 538;
Best Local Similarity 79.1%; Pred. No. 5.8e-46;
Matches 425; Conservative 0; Mismatches 106; Indels 6; Gaps 2;

QY 95 ttctcaatgtgtccctctctctcgtcggtgatatt--caataaaattcccttcgact 151
DB 1 TTCTCAATGTGTCCCTTCATCTCTGTGGCGATAGTGGCGAACTAAATTTCTTCCGATT 60
QY 152 gaggaactgattcccgagcattgtgtgtagacgcgagatagagctcgattgccagaaca 211
DB 61 GAGGCTGATCCCGACGATGTGTGTAAGACTGATATGAAGTTCGATTCGCGAATAAGCA 120
QY 212 aaccgtgttcaattacaatccagaattttcgacgtacaggaatttaactacagaagcta 271
DB 121 AACCATTTTCCATTACAGTCCGCGTAATTTCTACGTCACAGAGATTAACCTATACCAACT 180
QY 272 ctcaataagctacttgatccctcgtcctaataatgatcagaagaataattgcacagtttcc 331
DB 181 CTCAATCAGGCTACTCGATCCCGAGCTGAAATTCAGACCGAAATTCGCTGTGTTTCC 240
QY 332 aaatcacagggcaagttatgatgccaatgactagccaatcttgaatgggttcgtgttaa 391
DB 241 ACAACACAGGCAAGTACGATGCCATGACTAGCCCTAATCTTCGGATGGCTGTGTTGA 300
QY 392 caatgataatcaatagtcaactgtcctagctcctatcaattcgtcacagttatctctac 451
DB 301 TAACGATATCAATTATGTCAACTGTGCGGTCAAATCAACTCCTCAGTATACATTCCTAC 360
QY 452 aagttttttagcaaaaattcaa---cgggttttagctaaccttgcataagaagaatatt 508
DB 361 AAGTTTGTAGCACACAGCTACAAATTTCTAGTTTACGTATCTGTCTAGGTAATAATT 420
QY 509 gcaagcttcgatttggctggcggtgttagggttgaactgttgcattgtcctctgtctcc 568
DB 421 GCAAGCTTCAGATTTAGCAGTCGGTTGCAGAGTTGAACGTGTGGCATGGAGTCAGCTCC 480
QY 569 aggcatttcaacaacaagtcgtctacgttatacaagcacacatcaaggcctgctta 625
DB 481 GGGGATTTCACTTAACGTATCGTCTTCTGTTATCCAGCATTCACCAAGCTCTGGGTTA 537

RESULT 3
AI897599

LOCUS AI897599 533 bp mRNA linear EST 18-MAY-2001
DEFINITION EST267042 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
CLE28G17, mRNA sequence.
ACCESSION AI897599
VERSION AI897599.1 GI:5603501
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 533)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman,
C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley,
S.D. and Giovannoni, J.
TITLE Generation of ESTs from tomato carpel tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.
FEATURES
source
1..533
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLE28G17"
/clone_lib="tomato ovary, TAMU"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLED - Tomato Carpel EST Library. OligodT-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
BASE COUNT 135 a 103 c 120 g 175 t
ORIGIN
Query Match 24.9%; Score 301.6; DB 9; Length 533;
Best Local Similarity 75.3%; Pred. No. 2.1e-40;
Matches 405; Conservative 0; Mismatches 124; Indels 9; Gaps 2;
QY 305 tcagagagaattgcacagttttccaatcacagggcaagtatgatgccaactag 364
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 TCAGACCGAAATGCTCTGTTTTCACACACAGGCGCAAGTTACGATGCCATGACTAG 60
QY 365 ccaatcttgaatgggtcggtgtaacaatgatatacaactatgtaactgtctagctcc 424
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 CCTAATCTTCGGATGGCTTCGTGTGATTAACGATATCAATTATGTCAACTGTGCGGCTCA 120
QY 425 tatcaatcgtcacagtatattccataaagttttagtaaaaaattcaa--cggggtt 481
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 AATCAACTCTCAGTATACATCTCTCAAGAGTTTGTAGACACAGCTACAAATTCAGTTT 180
QY 482 tagctacctgtcataagaanaatatgcgaagcttcgaatttgctgagcgctgtagggt 541
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 181 TAGCTATCTGTGCTAGTGAATATATGCAAGCTTCAGATTAGCAGTCCGTTGCAGAGT 240
QY 542 tgaactgttgcatggttcctctgtccgaagcattcatcaaaacagtcgtctagttatc 601
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 241 TGAACCTGTGGCATGGAGTGCATCTCCGGGATTTACACCTAACGTATCGTCTGCTTATC 300
QY 602 aagcacacatcaaggcctggtctatggtttgagctttcttggaaagcgtaactgttatg 661
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 301 CAGCATTCACCAAGCTCTGGGTTACGGAATTGATCTCTCTTGAGAGGCGGTGATTTCTTATG 360
QY 662 tagaaattgcgacgcggagtcgtgggggtgagtgcaactattgaagaaaaacagcgacagagc 721

Db 361 CATAGAGTGTG-----GGAGAGATGATGATGTGTTTTCAAAGCAATAGTATGTAGC 414
QY 722 tacttgctgattatggtgcaaaagagacattcaagtttcgaagcttaagttccgatgcaa 781
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 415 AACTGTATTCATTACTGTAAAGAGGACACTCCTGTTTCTGAACGTTCTTCGGATGCAA 474
QY 782 agtcagttactattctgtttatgattgttctttggtgagtgataggaatagtgagtt 839
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 475 AGTCAGACTACTATTCGTTGTTTGTATCGATCTATGTTGGCATAGCAATTTGTGCACATT 532
RESULT 4
LOCUS AI490332 432 bp mRNA linear EST 18-MAY-2001
DEFINITION EST248658 tomato ovary, TAMU Lycopersicon esculentum cDNA clone
CLE24P14, mRNA sequence.
ACCESSION AI490332
VERSION AI490332.1 GI:4385642
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
Lycopersicon.
REFERENCE 1 (bases 1 to 432)
AUTHORS Alcalá, J., Vrebalov, J., White, R., Matern, A.L., Vision, T., Holt, I.E.,
Liang, F., Upton, J., Ronning, C.M., Craven, M.B., Fujii, C.Y., Bowman,
C.L., Nierman, W., Fraser, C.M., Venter, J.C., Martin, G.B., Tanksley,
S.D. and Giovannoni, J.
TITLE Generation of ESTs from tomato carpel tissue
JOURNAL Unpublished (1999)
COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>.
FEATURES
source
1..432
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLE24P14"
/clone_lib="tomato ovary, TAMU"
/tissue_type="carpel"
/dev_stage="5 days pre-anthesis to 5 days post-anthesis"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLED - Tomato Carpel EST Library. OligodT-primed and
directionally cloned cDNA in vector Lambda ZAP II with 5'
and 3' ends located at the EcoRI and XhoI sites,
respectively."
BASE COUNT 112 a 85 c 99 g 136 t
ORIGIN
Query Match 19.7%; Score 238.4; DB 9; Length 432;
Best Local Similarity 74.9%; Pred. No. 5.8e-30;
Matches 328; Conservative 0; Mismatches 101; Indels 9; Gaps 2;
QY 305 tcagagagaattgcacagttttccaatcacagggcaagtatgatgccaactag 364
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 1 TCAGACCGAAATGCTCTGTTTTCACACACAGGCGCAAGTTACGATGCCATGACTAG 60
QY 365 ccaatcttgaatgggtcggtgtaacaatgatatacaactatgtaactgtctagctcc 424
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 61 CCTAATCTTCGGATGGCTTCGTGTGATTAACGATATCAATTATGTCAACTGTGCGGCTCA 120
QY 425 tatcaatcgtcacagtatattccataaagttttagtaaaaaattcaa--cggggtt 481
||||| ||||||| ||||||| ||||||| ||||||| ||||||| ||||||| |||||||
Db 121 AATCAACTCTCAGTATACATCTCTCAAGAGTTTGTAGACACAGCTACAAATTCAGTTT 180

QY	482	tagctaacctgtcatcaagaagaataattgcaagcttcggaattggtgcgcgtgtgaagt	541
Db	181	TAGCTATCTTGTGCTAGGTGAATAATTGCAAGCTTCAGATTTAGCAGTCCGGTTGCAGAGT	240
QY	542	tgaactgttgcatagtctcctctgctccaggaattcatcaacaagtcgtctacgtatc	601
Db	241	TGAAACTGTGGCATGSACTGCAGCTCCGGGATTTCACTTAACGTATCGTCTTCGTATC	300
QY	602	aagcacacatcaagccttgcttatgggtttgagcttcttggaaagcgtaatctgtatg	661
Db	301	CAGCATTCACCAAGCTCTGGGTTACGATTTGATCTCTCTTGGAAGCGGTGATTTCTTATG	360
QY	662	tagaaattgcgaccggaatcgtgggggtgagtgcactattgaagaaacaagcgacacagc	721
Db	361	CAGAGAGTGTG-----GGAGAGATGATGATGTGTTCACAAGACATAGGGATGTAGC	414
QY	722	tacttgtcgttattgtgtg	739
Db	415	AACTGTATTCATTACTG	432

RESULT	5
BF645984	
LOCUS	BF645984
DEFINITION	NF065E11EC1F1086 Elicited cell culture Medicago truncatula CDNA clone NF065E11EC 5', mRNA sequence.
	654 bp mRNA linear EST 20-DEC-2000

ACCESSION	BF645984
VERSION	BF645984.1
KEYWORDS	GI:11911113
SOURCE	EST.
ORGANISM	barrel medic. Medicago truncatula

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae; Medicago.

REFERENCE	1 (bases 1 to 654)
AUTHORS	Torres-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation - Center for Medicago Genomics Research
JOURNAL	unpublished (2000)
COMMENT	Contact: Dixon RA

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380
Email: radixon@noble.org
Insert Length: 654 Std Error: 0.00
Plate: 065 row: E column: 11
Seq primer: TCACACAGGAACAAGCTATGAC.

```

FEATURES
    source
        1. .654
            location/Qualifiers
                /organism="Medicago truncatula"
                /db_xref="taxon:3880"
                /clone="NF065E11EC"
                /clone_lib="Elicited cell culture"
                /tissue_type="Cell cultures derived from root tissues"
                /dev_stage="Cell suspensions were subcultured every 14
                days. Cells were induced six days after subculture"
                /note="Vector: Lambda Zap; Cells were induced with yeast
                cell wall extracts equivalent to 50ug/ml glucose in the
                final concentration. Samples were taken at 0.5, 1, 12 and
                24 hours after induction. Equal amounts of RNA from each
                time point were pooled and used for mRNA isolation."
BASE COUNT
    214 a 145 c 102 g 192 t 1 others
ORIGIN

```

Query Match	5.8%;	Score 70;	DB 10;	Length 654;
Best Local Similarity	62.1%;	Pred. No. 0.026;		
Matches 128;	Conservative 0;	Mismatches 75;	Indels 3;	Gaps 1;

```

QY      102  tgtgtccctcttcctcgtcgatata--tcaataaaaattcccttcgactgagact 158
      ||| ||||| ||| ||| ||| ||||| ||||| ||||| |||
Db      165  TGTCCCCCTTCTTCATGTGGCAAAATCAGCAACATAAAGTATCCCTCCGACTTAAAGAT 224
QY      159  gatcccgagcatgtgtgtagacycgatatgagctcgattgccaagaacaaccaaccgtg 218
      ||||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      225  GATCCAGCAACCTGCGGTGATCCGAGGTACGAGTATTCCTCGCAAAAAAACATCACCATG 284
QY      219  ttcaattacaatocagaatlltcogacytacaggaatllaactacagaagctactcata 278
      ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      285  TTAAGCTGCTTCTTGTAATATCCATGTAAATCATCACTACATATACATACATAT 344
QY      279  aggcctactgatacctgycctaataatga 304
      ||| ||| ||||| ||| ||| |||
Db      345  CGGTAGTGTATCCTGGAAATAAAGA 370

```

RESULT	6				
BF633219					
LOCUS	BF633219	498 bp	mRNA	linear	EST 19-DEC-2000
DEFINITION	NF054H02DT1F1026 Drought	Medicago	truncatula	CDNA clone	NF054H02DT
	5', mRNA sequence.				
ACCESSION	BF633219				
VERSION	BF633219.1	GI:11897377			
KEYWORDS	EST.				
SOURCE	barrel medic.				
ORGANISM	Medicago truncatula				

REFERENCE	1 (bases 1 to 498)
AUTHORS	Torrez-Jerez, I., Scott, A.D., Harris, A.R., Gonzales, R.A., Bell, C.J., Flores, H.R., Inman, J.T., Weller, J.W. and May, G.D.
TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation Medicago truncatula drought library
JOURNAL	Unpublished (2000)
COMMENT	Contact: May GD

Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7391
Fax: 580 221 7380
Email: gdmay@noble.org
Insert Length: 498 Std Error: 0.00
Plate: 054 row: H column: 02
Seq primer: TCACACAGGAACACGCTATGAC.

FEATURES	
source	Location/Qualifiers
	1. .498
	/organism="Medicago truncatula"
	/db_xref="taxon:3880"
	/clone="NF054H02DT"
	/clone_lib="Drought"
	/tissue_type="Plantlets"
	/dev_stage="Pooled timepoints"
	/note="Vector: lambda zap; Contains a mixture of entire plantlets harvested in a series of days-post-watering timepoints."
BASE COUNT	167 a 110 c 75 g 146 t
ORIGIN	

Query Match	5.8%;	Score 69.8;	DB 10;	Length 498;
Best Local Similarity	59.2%;	Pred. No. 0.031;		
Matches 138;	Conservative	0;	Mismatches 92;	Indels 3; Gaps 1
QY 75	agttctaaagcgcgagatatattctcaatggtgcccttcttcctgcggtgatat---tcaa	131		
DB 169	ATTACTAATACAAATAACAAACATATGTGCCACCTTCTTCATGCGGCAAAATCAGCAAC	228		
QY 132	ataaaattcccttcgcactgaggactatcccgagcaattggtgtagaagcgagatatgag	191		
DB 229	ATCAACATCCCGTTTCGACTTAAGAAATGACCCAGCAACCCTGTGTGTGATCCAAAGTAGGAG	288		

[illegible]

RESULT	7				
AW695109					
LOCUS		603 bp	mRNA	linear	EST 21-DEC-2000
DEFINITION	AW695109				
	NE091D01ST1F1012	Developing stem	Medicago	truncatula	cdna clone
	NE091D01ST 5', mRNA sequence.				
ACCESSION	AW695109				
VERSION	AW695109.2	GI:11935567			
KEYWORDS	EST.				
SOURCE	barrel medic.				
ORGANISM	Medicago truncatula				

REFERENCE
AUTHORS
1 (bases 1 to 603)
He,X.-Z., Shadle,G., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell

TITLE	Expressed Sequence Tags from the Samuel Roberts Noble Foundation
JOURNAL	Medicago truncatula stem library
COMMENT	Unpublished (2000) On Apr 14, 2000 this sequence version replaced gi:7569871.

The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7302
Fax: 580 221 7380

Email: radixon@noble.org
Insert Length: 650 Std Error: 0.00
Plate: 091 row: D column: 01
Seq primer: TCACACAGGAACACGCTATGAC.

FEATURES	Location/Qualifiers
source	1. .603

```

/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF091D01ST"
/clone_lib="Developing stem"
/tissue_type="stem"
/dev_stage="Pooled developmental"
/note="Vector: Lambda Zap; Contains a mixture of
internodal stem segments"

```

Query Match	5.6%;	Score 67.8;	DB 9;	Length 603;
Best Local Similarity	61.7%;	Pred. No. 0.062;		
Matches 108; Conservative	0;	Mismatches 67;	Indels 0;	Gaps 0

QY 130 aaataaaattccctccgactgagactgatcccgaacatgtgtgtagcgcgatatg 189
| | | | | | | | | | | | | | | | | | | | |
Db 176 ACATAAACAATCCTTTTCGACTTCAGAATGATCCAGCAAACTCGGGTGATCCGAGGTATG 235

QY 190 agtcgcattggcagaacaaccaaacccgttgttcaattacaatccagaatttcgcagtac 249
| | | | | | | | | | | | | | | | | | | | |
Db 236 AGTATTCTTCGGAAAACACATCATCACACGTTAACTTGTTTTCAGSTAATACTATGTGA 295

QY 250 aggaattaactacagaagctactcaataaaggctacttgatcctcgycctaataga 304
| | | | | | | | | | | | | | | | | | | | |
Db 296 AATCAATCAACTACAGAAGTAAGTCAATGAGAGTAGTGTGATCTTGGAATTCAAGA 350

```

RESULT      8
LOCUS       AL513719/c
DEFINITION  AL513719 329 bp mRNA linear EST 13-FEB-2001
             AL513719 LTI_NFL005_PL2 Homo sapiens cDNA clone
             prime, mRNA sequence.
ACCESSION   AL513719
VERSION     AL513719.1 GI:12777213
KEYWORDS    EST.
SOURCE      human.
            Homo sapiens
ORGANISM    Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 329)
AUTHORS    Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE      Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
            location/Qualifiers
            1..329

```

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="CLOBA007ZH01"
/clone_lib="TRI_NFL006_PL2"
/tissue_type="placenta"
/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA
was primed with a NotI-oligo(dT) primer. Five prime end
enriched, double-stranded cDNA was digested with Not I and
cloned into the Not I and Eco RV sites of the pCMVSPORT 6
vector. Library was normalized. Library was constructed by
Life Technologies. Contact : Feng Liang Life Technologies,
a division of Invitrogen 9800 Medical Center Drive
Rockville, Maryland 20850, USA Fax : (1) 301 610 8371
Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com"

```

Query Match	5.6%;	Score 67.6;	DB 9;	Length 329;
Best Local Similarity	42.9%;	Pred. No. 0.085;		
Matches 73; Conservative	44;	Mismatches 53;	Indels 0;	Gaps 0

```

Qy      1041 ttttccttcagtttagtcctgatttattgatattcctacatttgattgtyrat 1100
        ||||| |||  :::: : | ||| ||| | ||| : ||| ||| |
Db      201 ttttttttttttKKKKKKKKKtttttttttttttttttttttttttttttt 142

```

[illegible]

QY	1161	tgtaatttctgtc	ataatagaca	agacattcg	caaaaaaaaa	1210
		: :: :: :	:	:	:	
Db	81	tttttttttttt	tttttttttt	tttttttttt	tttttttttt	32

RESULT	9
CNS03YE4/c	
LOCUS	
DEFINITION	CNS03YE4 1045 bp DNA linear GSS 18-MAY-2000 Tetraodon nigroviridis genome survey sequence T7 end of clone 068003 of library G from Tetraodon nigroviridis, genomic survey sequence.
ACCESSION	AL266197
VERSION	AL266197.1 GI:7987962
KEYWORDS	GSS; genome survey sequence.
SOURCE	Tetraodon nigroviridis.
ORGANISM	Tetraodon nigroviridis

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei; Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes; Tetraodontidae; Tetraodon.

REFERENCE 1 (bases 1 to 1045)
AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Fizames,C., Fisher,C.,
Bonneau,L., Billault,A., Quetier,F., Saurin,W., Bernot,A. and
Weissenbach,J.
TITLE Characterization and repeat analysis of the compact genome of the
freshwater pufferfish Tetraodon nigroviridis
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1045)
AUTHORS Roest-Crolius,H., Jallion,O., Dasilva,C., Bouneau,L., Fisher,C.,
Bernot,A., Fizames,C., Winkler,P., Brottier,P., Quetier,F.,
Saurin,W. and Weissenbach,J.
TITLE Human gene number estimate provided by genome wide analysis using
Tetraodon nigroviridis DNA sequence
JOURNAL Unpublished
REFERENCE 3 (bases 1 to 1045)
AUTHORS Genoscope.
TITLE Direct Submission
JOURNAL Submitted (12-APR-2000) to the EMBL/GenBank/DBJ databases
COMMENT This sequence is a single read and was generated as part of a large
scale clone-end sequencing project of the Tetraodon nigroviridis
genome. For more information, please take a look at
http://www.genoscope.cns.fr/Tetraodon.
FEATURES
source
1.1045
/organism="Tetraodon nigroviridis"
/db_xref="taxon:99883"
/clone="068003"
/clone_lib="G"
/note="Genoscope sequence ID : C0BG068AH02LP2-end : T7"
BASE COUNT 374 a 155 c 166 g 298 t 52 others
ORIGIN
Query Match 5.4%; Score 65.8; DB 12; Length 1045;
Best Local Similarity 50.3%; Pred. No. 0.11;
Matches 92; Conservative 24; Mismatches 67; Indels 0; Gaps 0;
QY 1028 ttaactagagattttcttcacagtttagtcctgtatttattgataattccta 1087
DB 955 TTTT TTTTAAATTTTAAATTTT TTTT TTTTAAATTTT TTTT TTTT 896
QY 1088 ttgattgtgtatgatttttcccttaaaatttataatttcctaattcctgta 1147
DB 895 TTTT TTTTANTTTT TTTT TTTTATTTT TTTT TTTTAAATTTT TTTT 836
QY 1148 ttgaatgatatgttacttctgtcacaataagaacaagacattcgcaaaaaa 1207
DB 835 TTAWAWMTTWT TTTT TTTTWT TTTT TTTTAAHAAARADARARAATTTT 776
QY 1208 aaa 1210
DB 775 AAA 773
RESULT 10
LOCUS BF637888 625 bp mRNA linear EST 19-DEC-2000
DEFINITION NF030B07PL1F1059 Phosphate starved leaf Medicago truncatula cDNA
clone NF030B07PL 5', mRNA sequence.
ACCESSION BF637888
VERSION BF637888.1 GI:11902046
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 625)
REFERENCE Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
AUTHORS 'H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula phosphate-starved leaf library

JOURNAL Unpublished (2000)
COMMENT Contact: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation
2510 Sam Noble Parkway, Ardmore, OK 73402, USA
Tel: 580 221 7325
Fax: 580 221 7380
Email: mharrison@noble.org
Insert length: 625 Std Error: 0.00
Plate: 030 row: B column: 07
Seq primer: TCACACAGGAACACGCTATGAC.
FEATURES
source
1.625
/organism="Medicago truncatula"
/db_xref="taxon:3880"
/clone="NF030B07PL"
/clone_lib="phosphate starved leaf"
/tissue_type="leaf"
/dev_stage="trifoliolate"
/note="Vector: Lambda Zap; At the trifoliolate stage, M.
truncatula plants were transplanted to phosphate-free sand
and grown for a further 30 days. During this 30 day
period, the plants were fertilized twice weekly with 1/2
Hoaglands solution containing only 20uM potassium
phosphate. RNA was prepared from above ground tissues."
BASE COUNT 206 a 134 c 101 g 184 t
ORIGIN
Query Match 5.4%; Score 65; DB 10; Length 625;
Best Local Similarity 57.9%; Pred. No. 0.18;
Matches 135; Conservative 0; Mismatches 95; Indels 3; Gaps 1;
QY 75 agttctaaagcgagatatcttcaatgtgtcccttctcctggtgatat--tcaa 131
DB 172 ATTACTAATACAAATAACCAACATATGTGCGACCTTCTTCATGCGGCAAAATCACA 231
QY 132 ataaattcccttcgagctgagcctgatacccgagcattgtgtagaacgagatatg 191
DB 232 ATATCACATCCTTTTCGACTTATGAAATGACCCCAACAGCTGTGTGATCCGAAGTACG 291
QY 192 ctgattgccagacaacacacccgtgttcaataacacacgaatttcgacgtacag 251
DB 292 TTATCTGCGGAACAACACACCGGTGTAACCTGTTTTCAGAGGTAATACTATGTGAC 351
QY 252 gaatttaactacagaagctactcaataaagctacttgatcctgacctaaatga 304
DB 352 GAAATCAACTACGTAATTTACACAATTGCGGCTAGTTGATCCAGTAATTGAAGA 404
RESULT 11
LOCUS BG455609 655 bp mRNA linear EST 19-MAR-2001
DEFINITION NF061D05PL1F1044 Phosphate starved leaf Medicago truncatula cDNA
clone NF061D05PL 5', mRNA sequence.
ACCESSION BG455609
VERSION BG455609.1 GI:13378934
KEYWORDS EST.
SOURCE barrel medic.
ORGANISM Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
Medicago.
1 (bases 1 to 655)
REFERENCE Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
AUTHORS 'H.R., Inman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
TITLE Expressed Sequence Tags from the Samuel Roberts Noble Foundation
Medicago truncatula phosphate-starved leaf library
JOURNAL Unpublished (2000)
COMMENT Contact: Harrison MJ
Plant Biology Division
The Samuel Roberts Noble Foundation

2510 Sam Noble Parkway, Ardmore, OK 73402, USA

Tel: 580 221 7325

Fax: 580 221 7380

Email: mjharrison@noble.org

Insert Length: 655 Std Error: 0.00

Plate: 061 row: D column: 05

Seq primer: TCACACAGGAAACAGCTATGAC.

FEATURES

source

1. .655
Location/Qualifiers

/organism="Medicago truncatula"

/db_xref="taxon:3880"

/clone="NF061D05PL"

/clone_lib="phosphate starved leaf"

/tissue_type="leaf"

/dev_stage="trifoliolate"

/note="Vector: Lambda Zap; At the trifoliolate stage, M.

truncatula plants were transplanted to phosphate-free sand

and grown for a further 30 days. During this 30 day

period, the plants were fertilized twice weekly with 1/2

Hoaglands solution containing only 20um potassium

phosphate. RNA was prepared from above ground tissues."

BASE COUNT

210 a 141 c 114 g 189 t 1 others

ORIGIN

Query Match

5.4%; Score 65; DB 10; Length 655;

Best Local Similarity 57.9%; Pred. No. 0.17;

Matches 135; Conservative 0; Mismatches 95; Indels 3; Gaps 1;

QY 75 agtctaaagcgagatattctcaatgtgtcccttccttcgtgatat---tcaa 131

Db 143 ATTACTAATACAAATTAACCAACATATGTCCACCTTCTTCATGCGCAAAATCACAAAC 202

QY 132 ataaattcccttcgactgagactgacccgagcattgtgtagacgcgatatgag 191

Db 203 ATATCACATCCCTTTTCGACTTATGAATGACCCACAGCTGTGTGATCCGAAGTACGAG 262

QY 192 ctgattgccagacaaccaacacggttcaattacaatccagaatttgcagctacag 251

Db 263 TTATCCTGCGAAAGACACATCACCGGTAACTCTGTTTCAGGTTAAATACTATGTCAG 322

QY 252 gaattactacagaagctactcaataagctacttgatcctgacctaatga 304

Db 323 GAAATCAACTAGCTAATTAATACACAATTCGGCTAGTTGATCCAGTAATTGAAGA 375

RESULT 12

BM358135

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

945 bp mRNA linear EST 09-JAN-2002
GA_Ea0005115r Gossypium arboreum 7-10 dpa fiber library Gossypium
arboreum cDNA clone GA_Ea0005115r, mRNA sequence.

BM358135
BM358135.1 GI:18098881
EST.

Gossypium arboreum.

Gossypium arboreum.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Rosidae; eurosids II; Malvales; Malvaceae; Gossypium.

1 (bases 1 to 945)

Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry

, D., Wood, T.C., Leslie, A. and Wilkins, T.A.

An integrated analysis of the genetics, development, and evolution

of the cotton fiber

Unpublished (2000)

Contact: Wing RA

Clemson University Genomics Institute

Clemson University

100 Jordan Hall, Clemson, SC 29634, USA

Tel: 864 656 7288

Fax: 864 656 4293

Email: rwing@clemson.edu

Total High Quality bases = 126

Seq primer: TAATACGACTCACTATAGCG

High quality sequence start: 3

High quality sequence stop: 842.

Location/Qualifiers

1. .945

/organism="Gossypium arboreum"

/strain="AKA"

/cultivar="8400"

/db_xref="taxon:29729"

/clone="GA_Ea0005115r"

/clone_lib="Gossypium arboreum 7-10 dpa fiber library"

/tissue_type="Fibers isolated from bolls harvested 7-10

dpa"

/lab_host="E. coli"

/note="Vector: pBK-CMV; Site_1: EcoRI; Site_2: XhoI"

484 a 57 c 18 g 368 t 18 others

BASE COUNT

484 a 57 c 18 g 368 t 18 others

ORIGIN

Query Match

5.4%; Score 64.8; DB 10; Length 945;

Best Local Similarity 52.1%; Pred. No. 0.16;

Matches 138; Conservative 0; Mismatches 127; Indels 0; Gaps 0;

QY 946 attttgtagtcaagctattttaacagaattgtgtatttttccagaaatctagga 1005

Db 634 ATTATTATTATATAAATATTGATAAAAGATAATTATAATTATTATTTT 575

QY 1006 caaggtcaacctgtgtcgcgattaattactagattttcttccagtttagtctgta 1065

Db 574 TTTTATTTTATTTTATTTTGTGTTTATTTTATTTTATTTTATTTTATTTT 515

QY 1066 ttatttgatatcttactatttgattgtgtatgatttttccctaaatttataa 1125

Db 514 TATTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTTATTTT 455

QY 1126 ttcttaattcttgtaagtaattgaaatgatatgttactttctgcaataatagaaca 1185

Db 454 TTTTATTAATTTTGTGTTAATGTTTATTTTATTTTATTTTATTTATTAATTTTCA 395

QY 1186 agacattcgcaaaaaa 1210

Db 394 TTTTATTTNNNAANNAANNAANNA 370

RESULT 13

AL513901/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

638 bp mRNA linear EST 13-FEB-2001
AL513901 LTI_NFL006_PL2 Homo sapiens cDNA clone CLOBA006ZG01 3
prime, mRNA sequence.

AL513901
AL513901.1 GI:12777395
EST.

human.

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

1 (bases 1 to 638)

Li, W.B., Gruber, C., Jessee, J. and Polayes, D.

Full-length cDNA libraries and normalization

Unpublished (2001)

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 Evry cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr.

Location/Qualifiers

1. .638

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="CLOBA006ZG01"

/clone_lib="LTI_NFL006_PL2"

/tissue_type="placenta"

/note="Vector: pCMVSPORT 6; Site_1: NotI; 1st strand cDNA

was primed with a NotI-oligo(dT) primer. Five prime end

enriched, double-stranded cDNA was digested with Not I and cloned into the Not I and Eco RV sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies. Contact : Feng Liang Life Technologies, a division of Invitrogen 9800 Medical Center Drive Rockville, Maryland 20850, USA Fax : (1) 301 610 8371 Email : fliang@lifetech.com URL : <http://fulllength.invitrogen.com>

	Query Match	5.3%;	Score 63.6;	DB 9;	Length 638;	
	Best Local Similarity	45.2%;	Pred. No. 0.3;			
	Matches 123;	Conservative 25;	Mismatches 124;	Indels 0;	Gaps 0;	
QY	939	tgtaagatttttctgtagtcaagcattttaacagaagtgtgtatatttttcagaaaa	998			
Db	331	TTTKKRGAAATTWAAAAAACCCTTTTTTTTTTAAATTTTTTTWTTTTTTTTTTWMMWT	272			
QY	999	tctagacaagftcaaccctgtgcgtggcgattaactaggatttttcttccagtttag	1058			
Db	271	ATTTAAARFTWTTTWTWTAWMTWTATAATWTTTCOCCTTTTTTTTTTTTTTTTWK	212			
QY	1059	tcctgtattttatttgatatcttacctatttgatgqtgatgattttttcccttaaat	1118			
Db	211	GTAATWATTTTTTTTTTTTTTTTTTTTWTTTTTTTTKKTATTTTKKWTTTTTTTTTTTT	152			
QY	1119	ttataaatlttccctaattctctgtaaglaattgaaatgsataattgtacttctgtccaataa	1178			
Db	151	TMMWTTTWWACCTTTTITTTTTRAGKTTATATAAAWMTAAAAATTITTATTTWAAA	92			
QY	1179	tagaacacacattcgcaaaaaaaaaaaaaa	1210			
Db	91	AAAAAAAAAAAAAAAAAAAAAAAAAAAAA	60			

RESULT	14			
CNS0012D				
LOCUS				
DEFINITION				
ACCESSION	CNS0012D	991 bp	DNA	linear
VERSION	Drosophila melanogaster genome survey sequence TERT3 end of BAC:			GSS 03-JUN-1999
KEYWORDS	BACR36013 of RPCI-98 library from Drosophila melanogaster (fruit fly), genomic survey sequence.			
SOURCE	AL074343			
ORGANISM	AL074343.1	GI:4954012		
	GSS.			
	fruit fly.			
	Drosophila melanogaster			
	Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;			
	Pterygota; Neoptera; Endopterygota; Diptera; Brachycera;			
	Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.			
REFERENCE	1 (bases 1 to 991)			
AUTHORS	Genoscope.			
TITLE	Direct Submission			
JOURNAL	Submitted (02-JUN-1999)	Genoscope - Centre National de Sequencage :		

COMMENT
<p>Determination of this BAC-end sequence was carried out as part of a collaboration with the Berkeley Drosophila Genome Project (BDGP). The BDGP is constructing a physical map of the Drosophila melanogaster genome using these BACs. For further information please see http://www.fruitfly.org The BDGP Drosophila melanogaster BAC library was prepared by Kazutoyo Osoegawa and Aaron Mammosser in Pieter de Jong's laboratory in the Department of Cancer Genetics at the Roswell Park Cancer Institute in Buffalo, NY. The library is named RPCI-98 and was constructed by partial EcoRI digestion of Drosophila DNA provided by the BDGP from the isogenic strain y2; cn bw sp, the same strain used for the BDGP's p1 and EST libraries. A more detailed description of the library and how to order individual BAC clones, the entire library, or filters for hybridization from the BACPAC Resource Center can be found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.</p>

Location/Qualifiers

```

source      1. .991
             /organism="Drosophila melanogaster"
             /db_xref="taxon:7227"
             /clone_lib="RPCI-98"
             /clone="BACR36013"
             /note="end : TET3"
BASE COUNT  176 a    95 c    64 g    515 t   141 others
ORIGIN

```

[illegible]

```

RESULT 15
LOCUS      AL515235
DEFINITION AL515235 LTI_NFL006_PL2 Homo sapiens CDNA clone Cl0BB021ZF02 3
ACCESSION  AL515235
VERSION     AL515235
KEYWORDS    prime, mRNA sequence.
SOURCE      AL515235.1  GI:12778728
            EST.
            human.
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.
REFERENCE   1 (bases 1 to 353)
AUTHORS    Li,W.B., Gruber,C., Jesse,J. and Polayes,D.
TITLE       Full-length cDNA libraries and normalization
JOURNAL     Unpublished (2001)
COMMENT     Contact: Genoscope
            Genoscope - Centre National de Sequencage
            BP 191 91006 EVRY cedex - France
            Email: segref@genoscope.cns.fr, Web : www.genoscope.cns.fr.
FEATURES
            source
            1..353

```

Query Match 5.2%; Score 62.6; DB 9; Length 353;

Best Local Similarity 54.4%; Pred. NO. 0.55;
Matches 92; Conservative 14; Mismatches 63; Indels 0; Gaps 0;

```
QY 1042 ttctctccagtttagtcctgattttatgtgataattcttacctattgattggtatg 1101
      ||| ||| ||| | | | | | | | | | | | | | | | | | | | | | | | |
Db 9 ttttttttttttttttttttttttttttttttttttttttttttttttttttttttt 68

QY 1102 atttttcccttaaaatttataattttccctaattctgtgaagtaattggaatgatatt 1161
      |||::| | | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 69 ttttwwtttttttttttttttttttttttttttttttttttttttttttttttttttt 128

QY 1162 gtacttctgtcaataatagacaagacatttcgcaaaaaaaaaaaaaaaaaa 1210
      : :: : | | | | | | | | | | | | | | | | | | | | | | | | | |
Db 129 TWTWWMWWMWMAAAAAAAAAAAAAAAAAAACACACAAAAAAAAAAAAAAAA 177
```

Search completed: June 4, 2002, 20:54:27
Job time: 1790 sec

OM of: US-09-805-919-1 to: N_Geneseq_032802:* out_format : pfs
Date: Jun 4, 2002 11:21 PM

About: Results were produced by the GenCore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh
-Q=/cgn2.1/USPTO.spool/US09805919/runat_04062002_121212_7676/app_query.fasta_1.369
-DB=N_Geneseq_032802 -QMT=fastap -SUFFIX=ring -GAPOP=12.000
-GAPEXT=4.000 -MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000
-GAPOP=4.500 -GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500
-FGAPOP=6.000 -FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500
-DELOP=6.000 -DELEXT=7.000 -START=1 -MATRIX=biosum62
-TRANS=human40.cdi -LIST=45 -DOCALLIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pfs
-NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000
-USER=US09805919_@CGN1.1.387 -NCPU=6 -ICPU=3 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-805-919-1
Query length: 308
Database: N_Geneseq_032802.*
Database sequences: 1736436
Database length: 858457221
Search time (sec): 203.120000

score_list:

Sequence	Strd Orig	ZScore	EScore	Len	1 Documentation
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:ABA01664 +		1642.00	3165.44	5.7e-16	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:ABA01665 +		1642.00	3162.37	8.5e-16	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA054855 +		128.50	221.04	0.0006	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA052806 +		127.50	219.07	0.0007	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA053268 +		108.00	181.86	0.0880	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL16690 +		105.50	162.45	1.06	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA054200 +		104.00	184.08	0.0662	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA052497 +		104.00	184.01	0.0668	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL26113 +		103.50	148.85	6.07	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL26112 +		103.50	143.29	12.38	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA039355 +		99.50	165.45	0.7216	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL16691 +		99.50	160.06	1.44	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA039374 +		94.50	141.62	15.34	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:AA039269 +		95.00	142.89	13.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:AA039250 +		95.00	142.89	13.04	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA075140 +		95.00	142.48	13.73	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA046755 +		94.50	143.88	11.48	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001A.DAT:AA094374 +		94.50	136.13	31.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:AA036035 +		94.50	136.13	31.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:AA026246 +		94.50	136.13	31.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1999.DAT:AA081825 +		94.50	136.13	31.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA088578 +		94.50	136.13	31.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA089796 +		94.50	136.13	31.03	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL06296 +		93.50	133.23	44.97	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA047184 +		93.00	144.84	10.15	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL04497 +		92.50	141.29	16.00	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1996.DAT:AA031288 +		92.00	155.22	2.68	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA079500 +		92.00	145.30	9.57	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:AA061796 +		92.00	145.30	9.57	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:AA051344 +		92.00	143.65	11.83	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA080223 +		92.00	143.65	11.83	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA080227 +		92.00	142.78	13.21	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA080229 +		92.00	141.96	14.69	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA080226 +		92.00	140.34	18.08	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA080233 +		92.00	139.40	20.39	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA082937 +		91.50	143.20	12.52	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1993.DAT:AA046025 +		91.50	139.34	20.55	
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1995.DAT:AA082938 +		91.50	139.34	20.55	

/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL06297 + 91.00 134.35 3
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL05109 + 91.00 120.39 2
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA1992.DAT:AA020995 + 90.50 144.61 10
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AA064276 + 90.50 139.72 19
/SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABL20591 + 90.50 111.45 7

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:ABA01664
seq_documentation_block:
ID ABA01664 standard; DNA; 927 BP.

XX ABA01664;
AC ABA01664;
DT 06-FEB-2002 (first entry)
XX
DE Nicotiana tabacum cv. Xanthinc stress induced gene SEQ ID NO:2.
XX
KW Nicotiana tabacum cv. Xanthinc; stress induced protein; plant;
KW stress induced gene; transgenic plant; resistance; ds.
XX
OS Nicotiana tabacum.

XX Key Location/Qualifiers
FT CDS 1..927
FT /tag= a
FT /product= "stress induced protein"
FT /trans_except= (pos:828..830,aa:Xaa)

PN JP2001252084-A.

PD 18-SEP-2001.

PF 15-MAR-2000; 2000JP-0071655.

PR 15-MAR-2000; 2000JP-0071655.

PA (NARA-) NARA SENTAN KAGAKU GIJUTSU DAIGAKUIN DAI.

DR WPI; 2002-029660/04.

DR P-PSDB; AAG68236.

PT New polypeptide for creating plant resistance to injury, comprises a
PT gene, the expression of which is induced by stress

PS Claim 3; Page 7-8; 11pp; Japanese.

XX The present sequence represents a gene isolated from tobacco, in which
CC the expression is induced by stress, and so encodes a stress induced
CC protein (I). The present invention also describes: (1) providing
CC resistance against stress by introducing the above gene to a plant; and
CC (2) a transformed plant which obtained resistance against stress by
CC introducing the above gene to the plant. The gene can be used for
CC creating plants resistant against injury, osmotic pressure, salt, or
CC low temperature-induced stress.

Sequence 927 BP; 256 A; 182 C; 212 G; 277 T; 0 other;

alignment_scores:
Quality: 1642.00 Length: 308
Ratio: 5.349 Gaps: 0
Percent Similarity: 99.675 Percent Identity: 99.675

alignment_block:
US-09-805-919-1 x ABA01664 ..

Align seg 1/1 to: ABA01664 from: 1 to: 927

1 MetLeuThrArgGlyLeuLeuPheAlaCysValLeuLeuValThrLe 17
|||||
1 ATGTTGACCAAGAGCGCTGCTTTTCGCTTGTGTGTTGTTGTTGACACT 50
17 uleserSerSerlyAlaGlnAspIleSerGlnCysValProSerSerC 34

|||||
51 CATAGCAGTCTTAAGCGCAGGATATTCTCAATGTGTCCCTTCTCCT 100
34 YSGIYASPILEGlnIleLysPheProPheArgLeuArgThrAspProGlu 50
101 GCGGTGATATTCAATAATAAATTTCCCTCCGACTGAGACTGATCCCGAG 150
51 HisCysGlyArgArgGlyTyrGluLeuAspCysGlnAsnAsnGlnThrVa 67
151 CATGTGTGAGACGCGGATATGAGCTCGATTGGCCAGAACCAACCAACCGT 200
67 lPheAsnTyrLysSerArgIlePheAspValGlnGluIleAsnTyrArgS 84
201 GTTCAATTACAATCCAGAATTTTCGACGTACAGGAATTAATACTACAGAA 250
84 eTyrSerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsn 100
251 GCTACTCAATAAGGCTACTGTATCCTCGCCCTAAATGATCAGAGAGAAAT 300
101 CysThrValPheProAsnHisArgAlaSerTyrAspAlaMetThrSergl 117
301 TGCACAGTTTTTCCAAATCACAGGGCAAGTTATGATGCCATGACTAGCCA 350
117 nIlePheGluTrpValArgValAsnAsnAspIleAsnTyrValAsnCysL 134
351 AATCTTGAATGGGTGCTGTGTTAACAATGATATCAACTATGTCAACTGTC 400
134 euAlaProIleAsnSerSerglTyrIleProThrSerPheCysSerLys 150
401 TAGCTCCTATCAATTCGTACACAGTATATCTCTACAAGTTTTTTGTAACAA 450
151 AsnSerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAlaSe 167
451 AATTCAACGGGTTTACCTACCTGTCTCAATAAGAGAAATATTGCAAGCTTC 500
167 rAspLeuAlaGlyGlyCysArgValGluThrValAlaTrpSerSeraLap 184
501 GGATTTGGCTGGCGGCTGTAGGGTTGAACCTGTTCATGTCTCTGCTC 550
184 roGlyIleSerSerAsnLysSerSerThrLeuSerSerThrHisGingly 200
551 CAGGCATTTTCATCAACAAGTCGCTACGTTATCAAGCACACATCAAGGC 600
201 LeuAlaTyrGlyPheGluLeuSerTrpLysArgAsnLeuLeuCysArgAs 217
601 CTGGCTTATGGGTTGAGCTTCTTGGAGCGTAATCTGTATGTAGAAA 650
217 nCysAspArgSerArgGlyGlyGluCysThrIleGlnGluAsnSeraSpa 234
651 TTGGACCGGAGTCGIGGGGGTGAGTGCACTATTGAAGAAAACAGCGACA 700
234 rGAlaThrCysArgTyrTrpCysLysGlnAspIleHisValSerLysLeu 250
701 GAGCTACTGTGCTTATGTGTCNAAGAGACATTCACGTTTCGAAGCTT 750
251 ThrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePheG1 267
751 ACGTCCGATGCAAAAGTCGAGTACTATTCGTTTATGTATTGTCTTGG 800
267 YGlyIleGlyIleGlyValLeuAla**ArgPheLeuLeuGlyIleP 284
801 CGGTATAGGAATAGGTGGGCTTTGGCGCTAAGATTCTACTAGAAATTC 850
284 roIleLeuIleAlaAlaValAlaTrpGlnCysLysArgArgAsnLeuHis 300
851 CAATCTTGATCGCAGCAGTGGTGTGCGAGTGCAAAAGACGAATTTGCAT 900
301 ThrSerSerAspGluGlnAsnCys 308
901 ACATCCTCCGATGAACAGAACTGT 924

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2002.DAT:ABA01665

seq_documentation_block:
ID ABA01665 standard; DNA; 1210 BP.
XX
AC ABA01665;
XX
DT 06-FEB-2002 (first entry)
XX
DE Nicotiana tabacum cv. Xanthinc stress induced gene SEQ ID NO:3.
XX
KW Nicotiana tabacum cv. Xanthinc; stress induced protein; plant;
KW stress induced gene; transgenic plant; resistance; ds.
XX
OS Nicotiana tabacum.
XX
FH Key Location/Qualifiers
FT CDS 18..943
FT /tag= a
FT /product= "stress induced protein"
FT /trans_except= (pos:846..848,aa:Xaa)
JP2001252084-A.
PD 18-SEP-2001.
XX
PF 15-MAR-2000; 2000JP-0071655.
XX
PR 15-MAR-2000; 2000JP-0071655.
XX
PA (NARA-) NARA SENTAN KAGAKU GIJUTSU DAIGAKUIN DAI.
XX
DR WPI; 2002-029660/04.
DR P-PSDB; AAG68236.
XX
PT New polypeptide for creating plant resistance to injury, comprises a
PT gene, the expression of which is induced by stress -
XX
PS Claim 4; Page 8; 11pp; Japanese.
XX
CC The present sequence represents a gene isolated from tobacco, in which
CC the expression is induced by stress, and so encodes a stress induced
CC protein (I). The present invention also describes: (1) providing
CC resistance against stress by introducing the above gene to a plant; and
CC (2) a transformed plant which obtained resistance against stress by
CC introducing the above gene to the plant. The gene can be used for
CC creating plants resistant against injury, osmotic pressure, salt, or
CC low temperature-induced stress.
SQ Sequence 1210 BP; 347 A; 217 C; 251 G; 395 T; 0 other;
XX
alignment_scores:
Quality: 1642.00 Length: 308
Ratio: 5.349 Gaps: 0
Percent Similarity: 99.675 Percent Identity: 99.675
alignment_block:
US-09-805-919-1 x ABA01665
Align seg 1/1 to: ABA01665 from: 1 to: 1210
1 MetLeuthArgGlyLeuLeuPheAlaCysValLeuLeuValThrle 17
|||||
18 ATGTTGACAGAGGGCTGCTTTCGCTTGTGTGTTGTTACTTGTGACACT 67
17 uIleSerSerSerLysAlaGlnAspIleSerglncCysValProSerSerc 34
|||||
68 CATAAGCAGTTCTAAAGCGCAGATATTCTCAATGTGTCCCTTCTTCT 117
34 YSGIYASPILEGlnIleLysPheProPheArgLeuArgThrAspProGlu 50
|||||
118 GCGGTGATATTCAATAATAAATTTCCCTCCGACTGAGAGACTGATCCCGAG 167
51 HisCysGlyArgArgGlyTyrGluLeuAspCysGlnAsnAsnGlnThrVa 67

|||||
168 CATGTGTGGTAGACGGCGATATGAGCTCGATGCGCCAGAACCAACCGT 217
67 lPheasNTyrLysSerArgIlePheAspValGlnGluIleAsnTyrArgS 84
218 GTTCAATTACAATCCAGAAATTTTCGACGCTACAGGAATTAACCTACAGAA 267
84 ertYrSerIleArgIleuAspProGlyLeuAsnAspGlnArgGluAsn 100
268 GCTACTCAATTAAGGCTACTTGATCCTGGCCCTAAATGATCAGAGAAAT 317
101 CysThrValPheProAsnHisArgAlaSerTyrAspAlaMetThrSerG 117
318 TGCACAGTTTTTCCAAATCACAGGGCAGCTTATGATGCCATGACTACCCA 367
117 nIlePheGluTrpValArgValAsnAsnAspIleAsnTyrValAsnCysL 134
368 AATCTTTGAATGGGTTCGTGTTAACAATGATATCAACTATGTCAACTGTC 417
134 euAlaProIleAsnSerSerGlnTyrIleProThrSerPheCysSerLys 150
418 TAGCTCCTATCAATTCGTACAGTATATCTTCTACAGTTTGTAGCATA 467
151 AsnSerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAlaSe 167
468 AATTCAACGGGTTTLAGCTACCTTGTCATAGAGAATAATTGCCAAGCTTC 517
167 rAspLeuAlaGlyGlyCysArgValGluThrValAlaTrpSerSerAlaP 184
518 GGATTTGGCTGGCGGCTGTAGGGTTGAACCTGTTGCATGGCTCTGCTC 567
184 roGlyIleSerSerAsnLysSerSerThrLeuSerSerThrHisGlnGly 200
568 CAGGCATTTTCATCAACAAGTCGTCTACGTTATCAAGCACACATCAAGGC 617
201 LeuAlaTyrGlyPheGluLeuSerTrpLysArgAsnLeuLeuCysArgAs 217
618 CTGGCTTATGGGTTGAGCTTCTTGGAAGCGTAATCTGTTATGTAAGAAA 667
217 nCysAspArgSerArgGlyGlyGluCysThrIleGlyGluAsnSerAspA 234
668 TTGGGACCGGAGTCGTGGGGGTGAGTGACATGTAAGAAACACAGCAGACA 717
234 rGAlaThrCysArgTyrTrpCysLysGluAspIleHisValSerLysLeu 250
718 GAGTACTTGTCTGTTATGGTGCATAAGAGGACATTCACGTTTCGAAGCTT 767
251 ThrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePheG 267
768 ACGTTCGATGCAAAAGTCGACTACTATCTGTTATGTAATGTTCTTGG 817
267 yGlyIleGlyIleGlyGlyValLeuAla**ArgPheLeuLeuGlyIleP 284
818 CGGTATAGGAATAGGTGAGTGGCGCTAAGATTCTACTAGGAATTC 867
284 roIleLeuIleAlaAlaValValTrpGlnCysLysArgArgAsnLeuHis 300
868 CAATCTTGATCGCAGCAGTGGTGGCAGTGCATAAAGACGGAATTTGCAT 917
301 ThrSerSerAspGluGlnAsnCys 308
918 ACATCCTCCGATGAACAGAACTGT 941

seq_name: /SIDsl/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT: AAC54855

seq_documentation_block:

ID AAC54855 standard; DNA; 1246 BP.

AC AAC54855;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 79315.

XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX Arabidopsis thaliana.
PN EP1033405-A2.
XX 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.

PR	23-JUN-1999;	99US-0140353.
PR	23-JUN-1999;	99US-0140354.
PR	24-JUN-1999;	99US-0140693.
PR	28-JUN-1999;	99US-0140825.
PR	29-JUN-1999;	99US-0140991.
PR	30-JUN-1999;	99US-0141287.
PR	01-JUL-1999;	99US-0141842.
PR	01-JUL-1999;	99US-0142154.
PR	02-JUL-1999;	99US-0142055.
PR	06-JUL-1999;	99US-0142390.
PR	08-JUL-1999;	99US-0142803.
PR	09-JUL-1999;	99US-0142920.
PR	12-JUL-1999;	99US-0142977.
PR	13-JUL-1999;	99US-0143542.
PR	14-JUL-1999;	99US-0143624.
PR	15-JUL-1999;	99US-0144005.
PR	16-JUL-1999;	99US-0144085.
PR	16-JUL-1999;	99US-0144086.
PR	19-JUL-1999;	99US-0144325.
PR	19-JUL-1999;	99US-0144331.
PR	19-JUL-1999;	99US-0144332.
PR	19-JUL-1999;	99US-0144333.
PR	19-JUL-1999;	99US-0144334.
PR	19-JUL-1999;	99US-0144335.
PR	20-JUL-1999;	99US-0144352.
PR	20-JUL-1999;	99US-0144632.
PR	20-JUL-1999;	99US-0144884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	22-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145224.
PR	26-JUL-1999;	99US-0145276.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	28-JUL-1999;	99US-0145919.
PR	02-AUG-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147038.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147260.
PR	06-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.

PR	31-AUG-1999;	99US-0151438;
PR	01-SEP-1999;	99US-0151930;
PR	07-SEP-1999;	99US-0155363;
PR	10-SEP-1999;	99US-0153070;
PR	13-SEP-1999;	99US-0155758;
PR	15-SEP-1999;	99US-0154018;
PR	16-SEP-1999;	99US-0154039;
PR	20-SEP-1999;	99US-0154779;
PR	22-SEP-1999;	99US-0155139;
PR	23-SEP-1999;	99US-0155486;
PR	24-SEP-1999;	99US-0155659;
PR	28-SEP-1999;	99US-0156458;
PR	29-SEP-1999;	99US-0156596;
PR	04-OCT-1999;	99US-0157117;
PR	05-OCT-1999;	99US-0157753;
PR	06-OCT-1999;	99US-0157865;
PR	07-OCT-1999;	99US-0158029;
PR	08-OCT-1999;	99US-0158232;
PR	12-OCT-1999;	99US-0158369;
PR	13-OCT-1999;	99US-0159293;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159294;
PR	13-OCT-1999;	99US-0159329;
PR	14-OCT-1999;	99US-0159331;
PR	14-OCT-1999;	99US-0159330;
PR	14-OCT-1999;	99US-0159637;
PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160817;
PR	21-OCT-1999;	99US-0160814;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161992;
PR	29-OCT-1999;	99US-0162142;

alignment_scores:		
Quality:	128.50	
Ratio:	1.179	
Percent Similarity:	52.913	
	Length:	206
	Gaps:	13
Percent Identity:	30.097	

alignment_block:
US-09-805-919-1 x AAC54855/rev ...

Align seg 1/1 to reverse of: AAC54855 from: 1 to: 1246

```

11 ValLeuLeuLeuValThrLeuIleSerSerSerLys.AlaglaSpiles 27
   :::::::::::::: :::::::::: :::::
1158 CTATTCTCTCTCTTCTCTTCTACTTTCCTCTCCTAAATGCCTCAAAACCA 1109

27 ergIncysValProSerSerCysGly.....AspIleGlnIleLysPhe 41
   :::::::::: :::::::::: :::::::::: ::::::::::
1108 AACGTGCTACAGTCTCTTCATGTGGAGGAACAACACCTGCAGCTTACGCTTC 1059

42 ProPheArgLeuArgThrAspProGluHis.....CysGlyArgArg 55
   :::::::::: ::::: :::::::::: :::::
1058 CATTTCGGCTA.....TCTCCTGAGCAATCTTCGTCATGTGACTACCC 1015

55 gGlyTrcIuLeuAspCysGlnAsnAsnGlnThrValPheAsnTrpLys 72
   :::::::::: ::::: :::::::::: :::::
1014 AGGATTCACCTCCATTGTGAGACAGCCTTAAAGCTTCCAAAT.....T 971

```

```
72 eArgIlePheAspValGlnGluIleAsnTyrArgSerTyrSerIleArg 88
      |||||  |||  |||||:::|  ::|||  ||||:::
970 CCAGACGATTTCTCGTCCAGACATA.....AAGTCGACGCGTATTCAC 927
      89 LeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPhePr 105
      |||  |||||  ::|||  ::|
926 CTTAGAGACCCC.....GATAACTGC.....TTAGC 901
      105 oAsnHisArgAlaSerTyrAspAlaMetThrSerGlnIlePheGluTrpV 122
      :::::  ||:::|||||  |||  ::|
900 CAGACGGCTTCTCAGCTTCGACGCTCAATGCTCTCTTTCTCTCTC 851
      122 aArgValAsnAsnAspIleAsnTyrValAsnCysLeuAlaProIleAsn 138
      |||||  :::::||||  :::::||||
850 ACCTTGTCAC.....TACACTTCTTGAGCTGTCAATACGAAACGTC 807
      139 SerSerGlnTyrIleProThrSerPheCysSerLysAsnSerThrglyPh 155
      ::|||:::  ::|  |||  |||||
806 AATCCTCTCTTGTGGAAACCATCCATTGCCCTTACTAATCCACA..... 762
      155 eSerTyrLeuValIleArgGluIleLeuGlnAlaSerAspLeuAlaGlyG 172
      761 .....CCTTCAGATCTCACCGGTT 743
      172 Ly.CysArg.....ValGluThrVa 178
      ::|||  :::::
742 CAATGCCGCTCTGTTCAAACCTTGAAGACACTACTTCTCCGGTCTCF 693
      178 LAleTTPSerSerAlaProGlyIleSerSerAsnLysSerSerThrLeus 195
      :  ||:::  |||||:::|||||:::|
692 TCACCACTCGCAGGTG.....ATCTCAACAACAAGACCTCTGGTGA 649
      195 eSerThrHisGln 199
      ::::|||||:::
648 ATGGACTCACCGG 635
```

seq_name: /sids1/gcgdata/hoid-geneseq/geneseqn-emb1/NA2000.DAT:AAC52806

seq_documentation_block:

ID AAC52806 standard; DNA; 1249 BP.

XX AAC52806;

DT 18-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 72505.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.

PR 25-MAR-1999; 99US-0126264.

PR 29-MAR-1999; 99US-0126785.

PR 01-APR-1999; 99US-0127462.

PR 06-APR-1999; 99US-0128234.

PR 08-APR-1999; 99US-0128714.

PR 16-APR-1999; 99US-0129845.

PR 19-APR-1999; 99US-0130077.

PR 21-APR-1999; 99US-0130449.

PR 23-APR-1999; 99US-0130510.

```
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
```

PR	20-JUL-1999;	99US-01448884.
PR	21-JUL-1999;	99US-0144814.
PR	21-JUL-1999;	99US-0145086.
PR	21-JUL-1999;	99US-0145088.
PR	22-JUL-1999;	99US-0145085.
PR	22-JUL-1999;	99US-0145087.
PR	22-JUL-1999;	99US-0145089.
PR	23-JUL-1999;	99US-0145145.
PR	23-JUL-1999;	99US-0145218.
PR	23-JUL-1999;	99US-0145226.
PR	26-JUL-1999;	99US-0145274.
PR	27-JUL-1999;	99US-0145913.
PR	27-JUL-1999;	99US-0145918.
PR	27-JUL-1999;	99US-0145919.
PR	28-JUL-1999;	99US-0146386.
PR	02-AUG-1999;	99US-0146388.
PR	02-AUG-1999;	99US-0146389.
PR	03-AUG-1999;	99US-0147203.
PR	04-AUG-1999;	99US-0147204.
PR	04-AUG-1999;	99US-0147302.
PR	05-AUG-1999;	99US-0147192.
PR	05-AUG-1999;	99US-0147303.
PR	06-AUG-1999;	99US-0147305.
PR	06-AUG-1999;	99US-0147416.
PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148171.
PR	11-AUG-1999;	99US-0148319.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154079.
PR	20-SEP-1999;	99US-0154739.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158363.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.

PR	14-OCT-1999;	99US-0159638;
PR	18-OCT-1999;	99US-0159584;
PR	21-OCT-1999;	99US-0160741;
PR	21-OCT-1999;	99US-0160767;
PR	21-OCT-1999;	99US-0160768;
PR	21-OCT-1999;	99US-0160770;
PR	21-OCT-1999;	99US-0160814;
PR	21-OCT-1999;	99US-0160815;
PR	22-OCT-1999;	99US-0160980;
PR	22-OCT-1999;	99US-0160981;
PR	22-OCT-1999;	99US-0160989;
PR	25-OCT-1999;	99US-0161404;
PR	25-OCT-1999;	99US-0161405;
PR	25-OCT-1999;	99US-0161406;
PR	26-OCT-1999;	99US-0161359;
PR	26-OCT-1999;	99US-0161360;
PR	26-OCT-1999;	99US-0161361;
PR	28-OCT-1999;	99US-0161920;
PR	28-OCT-1999;	99US-0161993;
PR	29-OCT-1999;	99US-0162142;

alignment_scores:		
Quality:	127.50	Length: 206
Ratio:	1.170	Gaps: 13
Percent Similarity:	52.913	Percent Identity: 29.612

alignment_block:

US-09-805-919-1 x AAC52806/rev ...

Align seg 1/1 to reverse of: AAC52806 from: 1 to: 1249

```

11 ValLeuLeuValuAlThrLeuIleSerSerSerLys.AlaGlnaspPiles 27
   :::::|:::::|  :::::||||||  ||  :::::|
1160 CFAATTCTCTCTCTCTTCTACTTTCTCTCTCTCCTAAATGCCTCAAAACCAA 1111
```

```

27  ergIncysvalProserSerCysGly.....AspilegInIleLysPhe 41
   ::::| | | | | | | | | | | | | | | | | | | | | |
110 AACGTGGCTACAGTTCTTCATGTGGAGSAAACACCTTGAAGTTAGCTTC 1061

```

42 ProPheArgLeuArgThrAspProGluHis.....CysGlyArg 55
 ||||| ||| ::|||::: |||||
 1060 CCCTTCTGGCTA....TCTCCTGACCAATCTCTGCTCTGGCGCTACC 1017

55 gcltyrgrlulcuaspcysglnasnasnglnrthValPheasntyrLys 72
 |||:::||||| |||::: :: |||
 1016 AGCATTCACCTCCATTGTGAGACAGCTTAAAGCTTCCAAT.....T 973

72 eArgilePheaspValngIngluileasntYrargSerTyrrSerllearg 88
 ||||| ||| |||||:::||||| ::||| |||:::
 972 CCAGACCATTTCTCGTCCAGACATT...AAATGCAGCGTATTCAC 929

89 LeuLeuaspProGlyLeuasnaspGlnargGluasnCysThrvalPhepr 105
 ||| |||||
 928 CTCAGAGACCCC.....GAATAACTGC.....TTAGC 903

105 oasnhisargalasertryaspalamethrserglnilepheglutryp 122
 : : : : : | | | | | : : :
 902 TAGCGGGCTTCTCAGGCTCGACGGCTTCAGGGTCTCTTTTCTCTCTTC 853

122 aIaRgValaIasnaSnaSPlIeasntYrValaIasncysLeuAlaProIleasn 138
:::|||||
852 ACCTGTGCAAC.....TACACTTCTTAAGCGTGCATCAGGAAACGTC 809
:::|||||

139 SerSerGlnTyrIleProThrSerPheCysSerIysAsnSerThrGlyPhe 155
808 AATTCCTCTCTTTGGAACTCATTCATTCGCTTGGCTAATTCACA 764
:::||||:: :: ||| |||||

155 eSerrYrLeuValIleargCluIleLeuGlInalaseRasplLeuAlaGlyc 172
763 |||||:||||: CCMTCAGATCTCTCAGCCAGCTT 745

172 ly.CysArg.....ValGIUThrVa 178


```

:: |||||
744 CAATGCCGCTCTTGTCAATCTTCAGACGCTACTTCTTCAGTCTAT 695
178 lAlaTrpSerSerAlaProGlylIleSerSerAsnlySerSerThrIeus 195
    : |||::
694 TCACCACCTCGCAGTGTG.....ACCTCAACAACAAGACCCTCTGGCTGAA 651
195 erSerThrHisGln 199
    : |||::
650 ATGGGACTCACCGG 637
```

seq_name: /SIDS1/gcgdate/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAC53268

seq_documentation_block:

ID AAC53268 standard; DNA; 1177 BP.

XX AAC53268;

DT 18-OCT-2000 (first entry)

XX Arabidopsis thaliana DNA fragment SEQ ID NO: 73938.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

OS Arabidopsis thaliana.

PN EP1033405-A2.

XX 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

```

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 07-MAY-1999; 99US-0132487.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
```

```

PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 26-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
```

PR	09-AUG-1999;	99US-0147493.
PR	09-AUG-1999;	99US-0147935.
PR	10-AUG-1999;	99US-0148317.
PR	11-AUG-1999;	99US-0148819.
PR	12-AUG-1999;	99US-0148341.
PR	13-AUG-1999;	99US-0148565.
PR	13-AUG-1999;	99US-0148684.
PR	16-AUG-1999;	99US-0149368.
PR	17-AUG-1999;	99US-0149175.
PR	18-AUG-1999;	99US-0149426.
PR	20-AUG-1999;	99US-0149722.
PR	20-AUG-1999;	99US-0149723.
PR	20-AUG-1999;	99US-0149929.
PR	23-AUG-1999;	99US-0149902.
PR	23-AUG-1999;	99US-0149930.
PR	25-AUG-1999;	99US-0150566.
PR	26-AUG-1999;	99US-0150884.
PR	27-AUG-1999;	99US-0151065.
PR	27-AUG-1999;	99US-0151066.
PR	27-AUG-1999;	99US-0151080.
PR	30-AUG-1999;	99US-0151303.
PR	31-AUG-1999;	99US-0151438.
PR	01-SEP-1999;	99US-0151930.
PR	07-SEP-1999;	99US-0152363.
PR	10-SEP-1999;	99US-0153070.
PR	13-SEP-1999;	99US-0153758.
PR	15-SEP-1999;	99US-0154018.
PR	16-SEP-1999;	99US-0154039.
PR	20-SEP-1999;	99US-0154779.
PR	22-SEP-1999;	99US-0155139.
PR	23-SEP-1999;	99US-0155486.
PR	24-SEP-1999;	99US-0155659.
PR	28-SEP-1999;	99US-0156458.
PR	29-SEP-1999;	99US-0156596.
PR	04-OCT-1999;	99US-0157117.
PR	05-OCT-1999;	99US-0157753.
PR	06-OCT-1999;	99US-0157865.
PR	07-OCT-1999;	99US-0158029.
PR	08-OCT-1999;	99US-0158232.
PR	12-OCT-1999;	99US-0158369.
PR	13-OCT-1999;	99US-0159293.
PR	13-OCT-1999;	99US-0159294.
PR	13-OCT-1999;	99US-0159295.
PR	14-OCT-1999;	99US-0159329.
PR	14-OCT-1999;	99US-0159330.
PR	14-OCT-1999;	99US-0159331.
PR	14-OCT-1999;	99US-0159637.
PR	14-OCT-1999;	99US-0159638.
PR	18-OCT-1999;	99US-0159584.
PR	21-OCT-1999;	99US-0160741.
PR	21-OCT-1999;	99US-0160767.
PR	21-OCT-1999;	99US-0160768.
PR	21-OCT-1999;	99US-0160770.
PR	21-OCT-1999;	99US-0160814.
PR	21-OCT-1999;	99US-0160815.
PR	22-OCT-1999;	99US-0160980.
PR	22-OCT-1999;	99US-0160981.
PR	22-OCT-1999;	99US-0160989.
PR	25-OCT-1999;	99US-0161404.
PR	25-OCT-1999;	99US-0161405.
PR	25-OCT-1999;	99US-0161406.
PR	26-OCT-1999;	99US-0161359.
PR	26-OCT-1999;	99US-0161360.
PR	26-OCT-1999;	99US-0161361.
PR	28-OCT-1999;	99US-0161920.
PR	28-OCT-1999;	99US-0161992.
PR	28-OCT-1999;	99US-0161993.
PR	29-OCT-1999;	99US-0162142.

```

alignment_scores:
  Quality: 108.00
  Ratio: 0.675
  Percent Similarity: 49.844
  Length: 321
  Gaps: 16
  Percent Identity: 22.741

```

```

alignment_block:
US-09-805-919-1 x AAC53268 ..
Align seg 1/1 to: AAC53268 from: 1 to: 1177

```

```

6  LeuLeuPheAlaCysValLeuLeuLeuValThrLeuIleSerSerly 22
   :::::::::: :::: ::::: ||| :::::
92  GTTAAATCTCTCCCTCCTCCTTATTTACCCCTCTCTGTGATTCAAG 141
   : ||| ::::: ||| ::::: ||| ::::: |||
22  sAlaGlnAspIleSerGlnCys..ValProSerSerCysGlyAspIleG 38
   : ||| ::::: ||| ::::: ||| ::::: |||
142 TAATCAAGGACTTGGCTGTGTGAGACTTTGTTTCAGTGTGAAACATCA 191
   : ||| ::::: ||| ::::: ||| ::::: |||
38  InIleLysPheProPheArgLeuArgThrAspProGlnHisCysGlyArg 54
   ||| ||| ||| ::::: ||| ::::: ||| ::::: |||
192 CCGGTGGTTTCCCTCTCTCGGGTGGGAATGCTCATAAAGATGCGGTCA 241
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
55  ArgGlyTyrGluLeuAspCys..GlnAsnAsnGlnThrValPheAsnTy 70
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
242 CCATCGCTGGAGCTTCACTGCAACAACAAACATCACCCTTTATTTCAT 291
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
70  rLysSerArgIlePheAspValGlnGluIleAsnTyrArgSerTyrSerI 87
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
292 CTCAAACCAAAAGTACTGTGTTCTCCGTATAGATCAAAACATCTAACAC 341
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
87  l earGLeuLeuAspProGlyLeu.....AsnAspGlnArg 98
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
342 TTACACTTGGCCAAACAGAACCTTCTAGGTCTTTTGTGCTCTGTATTC 391
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
99  GluAsnCysThrValPheProAsnHis.....ArgAlaSerTyrAs 112
   ||| ||| ::::: ||| ::::: ||| ::::: |||
392 ACAACACACAACCTTGCTCCCGAAACTTTCGAGCTTTCACGACCTACAA 441
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
112 PalamethrSerGlnIlePheGluTrpValArgValAsnAsnAspIleA 129
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
442 GAGCGTCACTATA...TTCTACCAATGCTCCTCTGTGCTTCCCTAACCT 488
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
129 snTyrValAsnCys.....LeuAlaProIleAsnSerSerGln..... 141
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
489 CTAGCTATACATGTCCTGAGATAGTCCCATCTCAGTGTCTGAAGCCCT 538
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
142 .....TyrIleProth 145
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
539 GAAAAATATCCTGAGAGCTGCTGTTCCAGTTTCACCCGGAAGGTTCCAC 588
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
145 rSerPheCysSerIys.....AsnSerThrGlyPheSerTyrL 158
   ||| ||| ::::: ||| ::::: ||| ::::: |||
589 GAGTTTCGATACAAAGAGAAAGAGTGAATGTGACCAATTGGAA... 634
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
158 euValIleArgGluIleLeuGlnAlaSerAspLeuAlaGlyGlyCysArg 174
   ||| ||| ||| ||| ::::: ||| ::::: ||| ::::: |||
635 .....AGTGTTTTAAGAAAAGGATTGTAG 658
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
175 ValGluThrValAlaTrpSerSerAlaProGlyIleSerSerAsnLys.. 190
   ||| ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
659 GTGACGGGTGTG.....ATCAATGAGAAATACATG 687
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
191 .SerSerThrLeuSerSerThrHisGlnGlyLeuAlaTyrGlyPheGluL 207
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
688 TCAGGAATGTTATCTCTCT...GGAAAGATGTCATGTCTTCAACG 731
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
207 euSerTrpLysArgAsnLeuLeuCysArgAsnCysAspArgSerArgGly 223
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
732 AAAACTTAACACCAAGAGTTAAATGTCTGCCACCAAGTAGCTGTAG... 778
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
224 GlyGlyCysThrIleGluGluAsnSerAspArgAlaThrCysArgTyrTr 240
   ||| ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
779 GGTTCCTGTGCATATAATCAGACCTCAAGTACGTTCCCTCTGC....TA 822
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
240 pCysLysGluAspIleHisValSerLysLeuThrPheArgCysLysValG 257
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||
823 TTGTAAGGATCCGTATAGCCTATCATGCAGTTCCGGCCAAATCTAAATTA 872
   ::::: ||| ||| ::::: ||| ::::: ||| ::::: |||

```


seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAC54200
seq_documentation_block:
ID AAC54200 standard; DNA; 494 BP.
XX
AC AAC54200;
XX
DT 18-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 77049.
XX
KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
KW metabolic pathway; promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-0301439.
XX
PR 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 04-MAY-1999; 99US-0132407.
PR 05-MAY-1999; 99US-0132484.
PR 06-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.

PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.

PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 13-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160776.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores: Quality: 104.00 Length: 110
 Ratio: 1.651 Gaps: 7
Percent Similarity: 57.273 Percent Identity: 31.818

alignment_block:
US-09-805-919-1 x AAC54200/rev ..

Align seg 1/1 to reverse of: AAC54200 from: 1 to: 494

1 MetLeuThrArgGlyLeuPheAlaCysValIleuLeuValThrIle 17
 :::|||||::: :::|||||
373 TTAAATATCATCTCTCTCTCTGTTTCTAATCTCTAATCCGACGACGAC 324

17 uIleSerSerSerLysAlaGlnAspIleSerGlnCysValProSerSerC 34
 ::: |||::: ::: |
323 GAGCACCACTTCACCGGT.....ACATGCACATAACGCCGTTT 286
34 ys.....GlyAspIleGlnIleLysPheProPheArgLeuArgThr 47
 ||| ||| |||:::||||| |||:::|
285 GTGCCCGTGCAGCGCCGATC...ATCCGTTTCCATTCGCTAAACAC 239
48 AspProGluHis...CysGly...ArgArgGlyTyrgIleuAspCysG 62
 ::: ||| ||| |||::: |||:::
238 CAACAATCATCTCTCTGTGTGTTACGACAAAGGCTTCGACTTAACGTGCGA 189
62 n.....AsnAsnGlnThrValPheAsnTyrlYsSerArgIleP 75
 |||::: |||::: |||
188 TATTAACGCCGGTAACAGAACACACCATTAACGTTACCGTTTCCGTAAC 139
75 heAspValGlnGluIleAsnTyArgSerTyrlSerIleArgLeuLeuAsp 91
 || |||:::||||| ::: ||| |||
138 TCACCGTGCAGAGATCGATTACGCACTCAAGAGATTGGATCAACGAC 89
92 ProGlyLeuAsnAspGlnArgGluAsnCys 101
 ||| :::|||||
88 CCA.....AACACTGT 77

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAC52497

seq_documentation_block:
ID AAC52497 standard; DNA; 497 BP.

XX AC AAC52497;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 71540.

XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway;
XX metabolic pathway; promoter; termination sequence; ss.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.
PR 05-MAR-1999; 99US-0123180.
PR 09-MAR-1999; 99US-0123548.
PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 23-APR-1999; 99US-0130891.
PR 28-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.

PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 14-JUN-1999; 99US-0138847.
PR 16-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.
PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.

PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 03-AUG-1999; 99US-0146389.
PR 04-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.
PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.

5237 ACTGCTTACCTGCCGTAGCAGGTGAGACTCAGCGGACCAACGTCATT 5286
159ValIleArgGlu.....IleuGlnA 166
5287 AGAAGTTAGCAGCGTCGAGGGGCTTTTCCGAGAAGAAGGGCTATTGCATT 5336
166 lAserAspleuAlacly.....GlyCysArgValGluThrValAlaTrp 180
5337 CGTGGGTTTCCCGCACTACCTGGCTGGCTGTGCAGTGGGTGCAGCC 5386
181 SerSerAlaProGlyIleSerSerAsnlySerSerThrIleuSerSerTh 197
5387 ACATCAGTGCCTGGTCGAGAGAGCTGCAGATCCTCATTGATTTCGGCCAC 5436
197 rHisGlnGlyLeuAlaTyrglyPheGluLeuSer..... 208
5437 TCCTTGGCCGCAAGCGTTGGAGCTTCAACTCCACCTGAGAGTTCCTTGT 5486
209TrpIlySArgAsnLeuLeuCysArgAsnCys 218
5487 GCAGCGCCAGCAATGCATTGACATCGTTCGAGATGCTGTGCAGCTCATCG 5536
219 AsparGSerArgGlyGlyGluCysThrIleGlu..... 229
5537 CAGGCGGCCAGGGGCATCGCTGGCTCATTCATGCGCTTGGCGCGCCTC 5586
230GluAsnSerAsparGAlaThrCysArgTyrrTrpCysL 242
5587 TGATTGCTCGCCATCGAAGTGCAGCTGTGCGAGCGATCGCATGTGTGCA 5636
242 ySgIuAspIleHisValSerIlySLeuThrPheArgCysIlyValGluTyr 258
5637 GCAC.....CACGCGCGCGGCGGCGTCTCGTCTGTCGCCGCCAC 5679
259 TyrSerValTyrValLeuPhePheGlyGlyIleGlyIleGlyValle 275
5680 TACAATGCAGTTGGCCAGCTCATCGTTGAGTT.....GGCAACAC 5720
275 uAla**ArgPheLeuLeuGlyIleProIleLeuIleAla.....Alav 290
5721 GCGGTTCACTTTCATCATCGTTAATGTTGTCTTCGTCACPTGATTGCCAA 5770
290 alValTrpGlnCysIlySArg 296
5771 TCACGTGCAGCTGTGCGCG 5790

seq_name: /SIDSI/gcgdata/hoid-geneseq/geneseqn-emb1/NA2000.DAT:AAC39355

seq_documentation_block:

ID AAC39355 standard; DNA; 1166 BP.

XX AAC39355;

DT 17-OCT-2000 (first entry)

DE Arabidopsis thaliana DNA fragment SEQ ID NO: 24330.

KW Hybridisation assay; genetic mapping; gene expression control;

KW protein identification; signal transduction pathway;

KW metabolic pathway; promoter; termination sequence; ss.

XX Arabidopsis thaliana.

OS Arabidopsis thaliana.

XX EP1033405-A2.

PD 06-SEP-2000.

PF 25-FEB-2000; 2000EP-0301439.

XX 25-FEB-1999; 99US-0121825.

PR 05-MAR-1999; 99US-0123180.

PR 09-MAR-1999; 99US-0123548.

PR 23-MAR-1999; 99US-0125788.
PR 25-MAR-1999; 99US-0126264.
PR 29-MAR-1999; 99US-0126785.
PR 01-APR-1999; 99US-0127462.
PR 06-APR-1999; 99US-0128234.
PR 08-APR-1999; 99US-0128714.
PR 16-APR-1999; 99US-0129845.
PR 19-APR-1999; 99US-0130077.
PR 21-APR-1999; 99US-0130449.
PR 23-APR-1999; 99US-0130510.
PR 28-APR-1999; 99US-0130891.
PR 30-APR-1999; 99US-0131449.
PR 30-APR-1999; 99US-0132048.
PR 30-APR-1999; 99US-0132407.
PR 04-MAY-1999; 99US-0132484.
PR 05-MAY-1999; 99US-0132485.
PR 06-MAY-1999; 99US-0132486.
PR 06-MAY-1999; 99US-0132487.
PR 07-MAY-1999; 99US-0132863.
PR 11-MAY-1999; 99US-0134256.
PR 14-MAY-1999; 99US-0134218.
PR 14-MAY-1999; 99US-0134219.
PR 14-MAY-1999; 99US-0134221.
PR 14-MAY-1999; 99US-0134370.
PR 18-MAY-1999; 99US-0134768.
PR 19-MAY-1999; 99US-0134941.
PR 20-MAY-1999; 99US-0135124.
PR 21-MAY-1999; 99US-0135353.
PR 24-MAY-1999; 99US-0135629.
PR 25-MAY-1999; 99US-0136021.
PR 27-MAY-1999; 99US-0136392.
PR 28-MAY-1999; 99US-0136782.
PR 01-JUN-1999; 99US-0137222.
PR 03-JUN-1999; 99US-0137528.
PR 04-JUN-1999; 99US-0137502.
PR 07-JUN-1999; 99US-0137724.
PR 08-JUN-1999; 99US-0138094.
PR 10-JUN-1999; 99US-0138540.
PR 10-JUN-1999; 99US-0138847.
PR 14-JUN-1999; 99US-0139119.
PR 16-JUN-1999; 99US-0139452.
PR 16-JUN-1999; 99US-0139453.
PR 17-JUN-1999; 99US-0139492.
PR 18-JUN-1999; 99US-0139454.
PR 18-JUN-1999; 99US-0139455.
PR 18-JUN-1999; 99US-0139456.
PR 18-JUN-1999; 99US-0139457.
PR 18-JUN-1999; 99US-0139458.
PR 18-JUN-1999; 99US-0139459.
PR 18-JUN-1999; 99US-0139460.
PR 18-JUN-1999; 99US-0139461.
PR 18-JUN-1999; 99US-0139462.
PR 18-JUN-1999; 99US-0139463.
PR 18-JUN-1999; 99US-0139750.
PR 18-JUN-1999; 99US-0139763.
PR 21-JUN-1999; 99US-0139817.
PR 22-JUN-1999; 99US-0139899.
PR 23-JUN-1999; 99US-0140353.
PR 23-JUN-1999; 99US-0140354.
PR 24-JUN-1999; 99US-0140695.
PR 28-JUN-1999; 99US-0140823.
PR 29-JUN-1999; 99US-0140991.
PR 30-JUN-1999; 99US-0141287.
PR 01-JUL-1999; 99US-0141842.
PR 01-JUL-1999; 99US-0142154.
PR 02-JUL-1999; 99US-0142055.
PR 06-JUL-1999; 99US-0142390.
PR 08-JUL-1999; 99US-0142803.
PR 09-JUL-1999; 99US-0142920.
PR 12-JUL-1999; 99US-0142977.
PR 13-JUL-1999; 99US-0143542.
PR 14-JUL-1999; 99US-0143624.
PR 15-JUL-1999; 99US-0144005.

PR 16-JUL-1999; 99US-0144085.
PR 16-JUL-1999; 99US-0144086.
PR 19-JUL-1999; 99US-0144325.
PR 19-JUL-1999; 99US-0144331.
PR 19-JUL-1999; 99US-0144332.
PR 19-JUL-1999; 99US-0144333.
PR 19-JUL-1999; 99US-0144334.
PR 19-JUL-1999; 99US-0144335.
PR 20-JUL-1999; 99US-0144352.
PR 20-JUL-1999; 99US-0144632.
PR 20-JUL-1999; 99US-0144884.
PR 21-JUL-1999; 99US-0144814.
PR 21-JUL-1999; 99US-0145086.
PR 21-JUL-1999; 99US-0145088.
PR 22-JUL-1999; 99US-0145085.
PR 22-JUL-1999; 99US-0145087.
PR 22-JUL-1999; 99US-0145089.
PR 22-JUL-1999; 99US-0145192.
PR 23-JUL-1999; 99US-0145145.
PR 23-JUL-1999; 99US-0145218.
PR 23-JUL-1999; 99US-0145224.
PR 26-JUL-1999; 99US-0145276.
PR 27-JUL-1999; 99US-0145913.
PR 27-JUL-1999; 99US-0145918.
PR 27-JUL-1999; 99US-0145919.
PR 28-JUL-1999; 99US-0145951.
PR 02-AUG-1999; 99US-0146386.
PR 02-AUG-1999; 99US-0146388.
PR 02-AUG-1999; 99US-0146389.
PR 03-AUG-1999; 99US-0147038.
PR 04-AUG-1999; 99US-0147204.
PR 04-AUG-1999; 99US-0147302.
PR 05-AUG-1999; 99US-0147192.
PR 05-AUG-1999; 99US-0147260.
PR 06-AUG-1999; 99US-0147303.
PR 06-AUG-1999; 99US-0147416.
PR 09-AUG-1999; 99US-0147493.
PR 09-AUG-1999; 99US-0147935.
PR 10-AUG-1999; 99US-0148171.
PR 11-AUG-1999; 99US-0148319.
PR 12-AUG-1999; 99US-0148341.
PR 13-AUG-1999; 99US-0148565.
PR 13-AUG-1999; 99US-0148684.
PR 16-AUG-1999; 99US-0149368.
PR 17-AUG-1999; 99US-0149175.
PR 18-AUG-1999; 99US-0149426.
PR 20-AUG-1999; 99US-0149722.
PR 20-AUG-1999; 99US-0149723.
PR 20-AUG-1999; 99US-0149929.
PR 23-AUG-1999; 99US-0149902.
PR 23-AUG-1999; 99US-0149930.
PR 25-AUG-1999; 99US-0150566.
PR 26-AUG-1999; 99US-0150884.
PR 27-AUG-1999; 99US-0151065.
PR 27-AUG-1999; 99US-0151066.
PR 27-AUG-1999; 99US-0151080.
PR 30-AUG-1999; 99US-0151303.
PR 31-AUG-1999; 99US-0151438.
PR 01-SEP-1999; 99US-0151930.
PR 07-SEP-1999; 99US-0152363.
PR 10-SEP-1999; 99US-0153070.
PR 13-SEP-1999; 99US-0153758.
PR 15-SEP-1999; 99US-0154018.
PR 16-SEP-1999; 99US-0154039.
PR 20-SEP-1999; 99US-0154779.
PR 22-SEP-1999; 99US-0155139.
PR 23-SEP-1999; 99US-0155486.
PR 24-SEP-1999; 99US-0155659.
PR 28-SEP-1999; 99US-0156458.
PR 29-SEP-1999; 99US-0156596.
PR 04-OCT-1999; 99US-0157117.
PR 05-OCT-1999; 99US-0157753.
PR 06-OCT-1999; 99US-0157865.

PR 07-OCT-1999; 99US-0158029.
PR 08-OCT-1999; 99US-0158232.
PR 12-OCT-1999; 99US-0158369.
PR 13-OCT-1999; 99US-0159293.
PR 13-OCT-1999; 99US-0159294.
PR 13-OCT-1999; 99US-0159295.
PR 14-OCT-1999; 99US-0159329.
PR 14-OCT-1999; 99US-0159330.
PR 14-OCT-1999; 99US-0159331.
PR 14-OCT-1999; 99US-0159637.
PR 14-OCT-1999; 99US-0159638.
PR 18-OCT-1999; 99US-0159584.
PR 21-OCT-1999; 99US-0160741.
PR 21-OCT-1999; 99US-0160767.
PR 21-OCT-1999; 99US-0160768.
PR 21-OCT-1999; 99US-0160770.
PR 21-OCT-1999; 99US-0160814.
PR 21-OCT-1999; 99US-0160815.
PR 22-OCT-1999; 99US-0160980.
PR 22-OCT-1999; 99US-0160981.
PR 22-OCT-1999; 99US-0160989.
PR 25-OCT-1999; 99US-0161404.
PR 25-OCT-1999; 99US-0161405.
PR 25-OCT-1999; 99US-0161406.
PR 26-OCT-1999; 99US-0161359.
PR 26-OCT-1999; 99US-0161360.
PR 26-OCT-1999; 99US-0161361.
PR 28-OCT-1999; 99US-0161920.
PR 28-OCT-1999; 99US-0161992.
PR 28-OCT-1999; 99US-0161993.
PR 29-OCT-1999; 99US-0162142.

alignment_scores: Quality: 99.50 Length: 321
 Ratio: 0.618 Gaps: 18
Percent Similarity: 50.156 Percent Identity: 21.495

alignment_block:
US-09-805-919-1 x AAC39355 ..

Align seg 1/1 to: AAC39355 from: 1 to: 1166

27 SerGlnCysValProSerSerCysGlyAspIleGlnIleLysPheProph 43
 |||||: :::|||||: |||:|:|:|:
176 TCACAGCCCTGCCAAAAACATGCGCCCAATACCCATCAAGTACCCGCT 225
 43 eArgLeuArgThrAspProGluHisCysGly.....ArgArgGlyTyrG 58
 : ::: ||||| |||
226 CGGAACCGGGTCAAGT.....TGCGGTGACCCACGCTTCACACGTT 266
 58 IuLeuAspCysGlnAsnAsnGlnThrValPheAsnTyrLysSerArgIle 74
 ::: |||:|:| :::|:|:| :::|:|
267 ACATCACGCTGACCCCTGACCAACAGACTTGACGCTCACCAACCCACACA 316
 75PheAspValGlnGluIleAsnTyrArgSerTyrSerIleArgIle 89
 ::: ::: :::|:|:|:|:|:| :::|:|:| :::
317 GGCAGTTACCCCTATCACTCTGTGCACTACGCCAACAAGATCTACGCT 366
 89 IuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPheProA 106
 : |||||:|:|:|:|:|:| |||:|:| |||:|
367 CACTGATCCGTCATGTCTACGTTGCA.....TGTACC...CGACCCA 407
 106 snHis.....ArgAlaSerTyrAspAlaMetThrSer 116
 ::|:| ||| :::|:|:| :::
408 GCCATGATTCGGCTTGATTTGGACGCTCCGTTTCTTTCATGACGAC 457
 117 GlnIlePheGluTrpValArgValAsnAsnAspIleAsnTyrValAsnC 133
 ::|:|:| ::: ||| :::|:| |||
458 ACCGTTTACGCTCCTTGACTGTCCGTTAGACGAATCCCTGTA..... 502
 133 sleuAlaProIle...AsnSerSerGlnTyrIleProThrSerPheCys 149
 :::|:|:|:|:|:| |||:|:|:|:| ::: |||:|:|:|:|:|
503 .TTCACACCGTTAAGTAACGGTAGCGACGCTGT.....TCCCTCTGTG 545

149 erlyAsnSerThrGly.....PheSerTyrLeuValIleArgGlu 162
:::|||||::: ||| |||
546 ATCGTCAAGCTCCCTCCATTTCACCTTCTTGTACTCAAAATTGCAGGGCC 595
163 IleLeuGlnAlaSerAspLeuAlaGlyGlyCysArgValGluThrValAl 179
||| ::| ::| ::| |||
596 ATAGCCTCATAAACCTCCAGATATCAACGTGTGCGTC..TACGTGCC 642
179 atpSerSerAlaProGlyIleSerSerAsnLysSerSerThrLysers 196
::: ||| ::| ::| ::|
643 TCTAGACCTCGGACCATCGTTCGAAATGGATTGAAACAACCTCAAGTGCT 692
196 ertHrHisGlnGlyLeu..... 201
||| ::| ::| ::|
693 CGTCATACTCCGGCTTCACAACTCGGTCCAGCCCAAGAAATCCACCCG 742
202AlaTyrGlyPheGluLeuSerTrpLysArgAsnLeu..... 213
||| ||| ||| |||
743 GAGAAATTGGAATACGAGATCGCACTCAAGTACAAAGTTCAACGTTTGA 792
214LeuCysArgAsnCysAspArgSerArgGlyGlyGluC 226
::| ::| ::| ::| ::|
793 TGAGTACCTCGCGTCTCGGTAGCTGCGAGCGTAGTAACGTGCT...T 839
226 ystHrIleGluGluAsnSerAspArgAlaThrCysArgTyrTrpCysLys 242
||| ::| ::| ::| ::|
840 GTGGCTTCAACACACAGTCAAGCTCGTTCGTGTGTAAAC.....TGTCCC 883
243 GluAspIleHisValSerLysLeuThrPhe.....Ar 253
||| ::| ::| ::|
884 GGTGGGATTAAACACGACATCGGATGCTCTCTCTTATTAACCTGTCTC 933
253 gCysLysValGluTyrTyrSer.....ValTyrValLeuPheP 266
||| ||| ||| |||
934 TATTCTTGTTCCTTGGTTCACAGAGATGGTAATGTACCCCTTGTGAT 983
266 heGlyGlyIleGlyIleGlyGlyValLeuAla***ArgPheLeuLeuGly 282
::| ::| ::| ||| ||| |||
984 GGTGTGTGTGTTGGATGGTGTGTTCTCGCTTAAGAGAGCATCAAGGA 1033
283 IleProIleLeuIleAlaValAlValTrpGlnCysLysArgArgAsnLe 299
||| ::| ::| ::| |||
1034TGGAGAGGTTCAAAAGAGGAAGCT 1056
299 uHisThrSerSer 303
||| ::| ::| ::|
1057 TAATGATCGAAT 1069

seq_name: /SIDS1/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:ABLI6691

seq_documentation_block:

ID ABLI6691 standard; DNA; 1863 BP.

XX ABLI6691;
AC
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 1546.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US09231.
XX
PR 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.

XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signalling and cell-cell
PT interactions -
XX
PS Claim 1; SEQ ID NO 1546; 21pp + sequence listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences and pharmaceutical drugs. The invention
CC sequences (ABLI01840-ABLI6175) and the encoded proteins
CC (ABB57737-ABB72072).
CC The sequence data for this patent did not form part of the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX
SQ Sequence 1863 BP; 419 A; 496 C; 486 G; 462 T; 0 other;

alignment_scores:
Quality: 99.50 Length: 268
Ratio: 0.816 Gaps: 13
Percent Similarity: 45.522 Percent Identity: 21.642

alignment_block:
us-09-805-919-1 x ABLI6691 ..

Align seg 1/1 to: ABLI6691 from: 1 to: 1863

27 SerGlnCysValProSerSerCysGlyAspIleGlnIleLysPheProPh 43
||| ::| ::| ::| ::| |||
39 TCTGAGTGTTCGCCAACAACCTGTTGAAACGCAATCCGAG..... 80
43 eArgLeuArgThrAspProGluHisCysGlyArg.ArgGlyTyrGluLeu 59
||| ::| ::| ::| ::| ||| |||
81CGAGGGCAACCGAATCAGGCCGCGAGAGCGAGGTACACAGGAA 123
60 AspCysGlnAsnAsnGlnThrValPheAsnTyrLysSerArgIlePheAs 76
||| ::| ::| ::| ::|
124 GACGGCGCAACGAGG..... 138
76 pValGlnGluIleAsnTyrArgSerTyrSerIleArgLeuLeuAspProG 93
||| ::| ::| ::| ::|
139TCACCAC 145
93 lYleuAsnAspGlnArgGluAsnCysThrValPheProAsnHisArgAla 109
::| ::| ::| ::|
146 CTGCACAAAGACACA.....ATGGGC 165
110 SerTyrAspAlaMetThrSerGlnIlePheGluTrpValArgValAsnAs 126
||| ::| ::| ::| ::|
166 AGCAGCGCCACGCAACCAACGAGTCCGTCGTGCAATT..... 204
126 nasPileAsnTyrValAsnCysLeuAlaProIleAsnSerSerGlnTyrI 143
||| ||| ||| |||
205TGTTATGTCTGTGATCGCCACGCTGTACG 235
143 leProThrSerPheCysSerLysAsnSerThrGlyPheSerTyrLeuVal 159
::| ||| ::| ::| ::|
236 TGTGCGTCTCCAGCGAGATGAGCACAAT....CACAGTACAAATGAC 279
160 IleArgGluIleLeuGlnAlaSerAspLeuAlaGlyGlyCys.....Ar 174
::| ||| ::| ::| ::|
280 CGGAGGAGGTGTGACTGTGATGAACACGCTGGCCCGGTACACAAATCG 329

```
174 gvaIgluThrvalaIatrpSerSerAlaProGlyIleSerSerAsnLysS 191
      ||||| |||: ||| ||| |||: |||
330 GCAGGAGACGTACCGCTACTTCTCTCTCCCTTTTGAGTGGCCAGAAGT 379
      |||||
191 erSer.....ThrLeuSerSerThrHisGlnGlyLeu 201
      ||||| |||: ||| ||| |||: |||
380 CCTCGATATCGCACTACCAAGAGACCGCTGAGCGGCTGCAAGGAGTTC 429
      |||||
202 AlaTyr.....GlyPheGluLeuSerTrpLys.....Ar 211
      |||: |||: |||: |||: |||: |||
430 GAGCTAGAGTTACGTGGCTACGAGATGAGTTCAAGAGCGAGCCCCCAA 479
      |||: |||: |||: |||: |||: |||
211 gasnLeuLeuCysArqasnCysAspArgSerArgGlyGlyGlyCysThri 228
      |||: |||: |||: |||: |||: |||
480 ATCGGTCACTGTGATG.....GTCACT 502

228 legluGluAsnSerAspArgAlaThrCysArgTyrTrpCysLysGluAsp 244
      |||: |||: |||: |||: |||: |||
503 TGCAGGAGAGAGCGCCAAAGCA..... 525

245 ILeHisValSerLysLeuThrPheArgCysLysValGluTyrTrpSerVa 261
      |||: |||: |||: |||: |||: |||
526 .....TTCACTATGCGGTGAAGAAGAGTACTG..... 555

261 lTyrvAlleuPhePheGlyIleGlyIleGlyIleGlyValLeuAla**A 278
      |||: |||: |||: |||: |||: |||
556 .TACCAGATGTACATGATGGACTGCCCATTTGGGGAAGAGTCGTGAGC 604

278 rg 278
605 GC 606
```

seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2000.DAT:AAA79720

seq_documentation_block:

ID AAA79720 standard; cDNA; 2040 BP.

AC AAA79720;

DT 27-NOV-2000 (first entry)

DE Eucalyptus grandis cell signalling involved polynucleotide SEQ ID NO:872.

KM Eucalyptus grandis; Pinus radiata; Monterey pine; plant; modification;
KW plant cell signalling; modulation; transgenic plant; pathogen; growth;
KW environmental change; development; cell proliferation; differentiation;
KM elongation; survival; disease resistance; nutrient metabolism; ss.

OS Eucalyptus grandis.

PN WO200042171-A1.

PD 20-JUL-2000.

PF 11-JAN-2000; 2000WO-US00724.

PR 12-JAN-1999; 99US-0228986.

PR 01-NOV-1999; 99US-0162866.

PA (GENE-) GENESIS RES & DEV CORP LTD.

PI Strabala TJ, Nieuwenhuizen NJ;

DR WPI; 2000-476052/41.

PT Isolated polynucleotide encoding a polypeptide involved in cell
PT signaling used for generating transgenic plants with modified responses
PT to external signals -

PS Claim 1; Page 427; 527pp; English.

CC AAA79263 to AAA79736 and AAB25100 to AAB25570 represent polynucleotide

CC and protein sequences isolated from eucalyptus (Eucalyptus grandis) or
CC pine (Pinus radiata also known as Monterey pine). The protein sequences
CC are involved in cell signalling. The polynucleotide and protein
CC sequences can be used to modify the response of plant cells to external
CC signals e.g. environmental changes or pathogens during the growth and
CC development of a plant. They can be used to modify cell proliferation,
CC differentiation, elongation and survival, resistance to disease and
CC nutrient metabolism. Examples of modifications which can be produced are
CC altered fruit ripening and senescence of leaves and flowers e.g. to
CC delay senescence and prolong the life of cut flowers or enhance
CC senescence of reproductive organs to engineer sterile plants. Other
CC modifications can be used to delay senescence in selected cell types or
CC organs providing fruit and vegetables which have a longer shelf life
CC between harvest and consumption, or to decrease branching frequency in
CC forest tree species giving long stretches of valuable knot-free clear
CC wood which can be used in solid timber furniture and veneers.

XX Sequence 2040 BP; 511 A; 522 C; 544 G; 463 T; 0 other;

alignment_scores:

Quality:	95.00	Length:	247
Ratio:	0.748	Gaps:	16
Percent Similarity:	51.417	Percent Identity:	23.482

alignment_block:

US-09-805-919-1 x AAA79720 ..

Align seg 1/1 to: AAA79720 from: 1 to: 2040

```
6 LeuLeuPheAlaCysValLeuLeuValThrLeuIleSerSerSerly 22
      ||||| |||
65 CTCCTCTCTCTCTCTC.....GCCATGGCACCGCGCA 99

22 sAlaGlnAspIleSerGlnCysValProSerSerCys.....G 35
      ||| ||| |||: |||
100 CCCCACTGACAAAGGCCCTTGCCGCCCCAGAGGTGCGCGCCCATCAGC 149

35 lYAspIleGlnIleLysPheProPheArgLeuArgThrAspProGluHis 51
      |||: |||: |||: |||: |||: |||
150 GGCGGTG...GTCCAGTTCCTGCTGCTGCAAGCAAGGCCAACAT 196

52 .....CysGly...ArgArgGlyTyrGluLeuAspCys...GlnAsnAs 64
      ||||| ||| |||: |||: |||: |||
197 GAGGACTGGCGCTACAGTCGGGACTTCACCTGTCTCGACACAGAGCGG 246

64 nGlnThrValPheAsnTyrLysSer.....ArgIlePheAspValGlnG 79
      |||: |||: |||: |||: |||: |||
247 GAACACCCCTGCTCTGCTCCCGGGGACCCCGCGGCACCTTGCTCACGG 296

79 lulLeAsnTyrArgSerTyrSerIleArgLeuLeuAspProGlyLeuAsn 95
      |||: |||: |||: |||: |||: |||
297 ACATCGATACGAGGCGAGTTCATCAGCGTGCGGGAC..... 334

96 AspGlnArgGluAsnCysThrValPhePro.....AsnHisAr 108
      |||: |||: |||: |||: |||: |||
335 .....CGGACGGCTGCATCCCTTGCCGCTGGCGAACCTCAACCACTC 378

108 gAla...SerTyrAspAlaMetThrSerGlnIlePheGluTrpValArgV 124
      ||||| |||: |||: |||: |||: |||: |||
379 TATCTCTCTCTACAG.....AAGACCAACCCAGCCAGTATAGTACT 422

124 alAsnAsnAspIleAsnTyrValAsnCysLeuAlaProIleAsnSer... 139
      |||: |||: |||: |||: |||: |||
423 ACCAGCGCGCCCTACAACATATACCTGTTCAGTGTCCGGGGAGGAGG 472

140 SerGlnTyrIleProThrSerPheCysSerLysAsnSerThrGlyPheSe 156
      |||: |||: |||: |||: |||: |||
473 AGCGAGGGCCGATGACCCGGGTGATTCACAGACACATCTGCGTATGA 522

156 rTyrLeuValIleArgGluIleLeuGlnAlaSerAspLeuAlaGlyGlyC 173
      |||: |||: |||: |||: |||: |||
523 CGTCTTCGCCCTCGAGTCTCTACTCTACGTGGAGACTTG..... 562
```


173 ysargValGluThrValAlaTrpSerSerAlaProGlyIleSerSerAsn 189
563
190 LysSerSerThrLeuSerSerThrHisGlnGlyLeuAlaTyrgIyPheG1 206
581 AGAGCGGAGCATGCTGCTCCCGAGATGATTAGTATGATGGCGGA 630
206 u.....LeuSerTrpLysArgAsnLeuLeuCysArgAsnCysAspArgS 221
631 GGTTAATTGGTCTGGGAAAAGCCG...GACTGCAGGCTATGCACGAAGA 677
221 erArgGlyGlyGlyCysThrIleGluGluAsnSerAspArg 234
678 AC...GGTCTGTGTCTGGGGAGACGAGACGGCATCCGC 715

seq_name: /SIDSL/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:AAS92969

seq_documentation_block:

ID AAS92969 standard; cDNA; 3874 BP.
XX AAS92969;
AC
XX
DT 13-FEB-2002 (first entry)
XX
DE DNA encoding novel human diagnostic protein #28773.
XX
KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX
OS Homo sapiens.
XX WO200175067-A2.
PN
XX
PD 11-OCT-2001.
XX
PF 30-MAR-2001; 2001WO-US08631.
XX
PR 31-MAR-2000; 2000US-0540217.
PR 23-AUG-2000; 2000US-0649167.
XX
PA (HYSE-) HYSEQ INC.
XX
PI Drmanac RT, Liu C, Tang YT;
XX
DR WPI: 2001-639362/73.
DR P-PSDB; ABG28782.
XX
PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity
XX
PS Claim 1; SEQ ID No 28773; 103pp; English.
XX
CC The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags
CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AAS64197-AAS94564 represent novel human

CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.
XX

SQ Sequence 3874 BP; 824 A; 1213 C; 1059 G; 778 T; 0 other;

alignment_scores:

Quality:	95.00	Length:	341
Ratio:	0.605	Gaps:	20
Percent Similarity:	46.041	Percent Identity:	20.821

alignment_block:

US-09-805-919-1 x AAS92969 ..

Align seg 1/1 to: AAS92969 from: 1 to: 3874

29 CysValPro.....SerSerCysGlyAspIleGlnIleLysPhe 41
2495 TGGTTCCTGGGAAGGGCTGCGAGGCTGC..... 2524
41 eProPheArgLeuArgThrAspProGluHisCysGlyArgArgIyTyrg 58
2525 .CCTCCAACCTGACTGGGGATGGCGGACACTGTGCGGCTCTGGGAGCT 2573
58 luleuAspCysGlnAsnAsnGlnThrValPheAsnTyr..... 70
2574 CTTCTCTGTGTGAGAACAGACAGTCTGCCCTGTGAATTACTGTACAAATCAA 2623
70 70
2624 GGCCACTGTACATCTCCAGACTCTGGCTGTACGCCATGTGCACCTG 2673
71LysSerArgIlePheAspValGlnIleAsnT 82
2674 CCCCCAGCCCTTCACTGACAGCGCTGCTTC.....CTGGCTG 2711
82 yfArgSerTyrSer.....IleArgLeuLeu 90
2712 GGAACAACCTCAGTCCCACTGTCAACCTAGAACTTCCCTTAAGATCATC 2761
91 AspProGlyLeuAsnAspGlnArgGluAsnCysThrValPheProAsnH1 107
2762 CAGCTCTGTGCTCAGTGAAGAGAAATGCCCTCCATGGCAGAGTCAACGC 2811
107 sArgAlaSerTyrAspAlaMetThrSerGlnIlePheGluTrpValArgV 124
2812 CTGGGTGGCATACAGACTGGGGAACCTGGACATGCGGCGCTTCTCCGC. 2860
124 alaAsnAsnAspIleAsnTyrValAsnCysLeuAlaProIleAsnSerSer 140
2861 ..AACAGCCCAAGTGAACGAATCATCTGCAGCACCGCGCTCGGAAGC 2908
141 GlnTyrIlePro.ThrSerPheCysSerLysAsnSerThrGlyPheSerT 157
2909 CCC..ATCCAACACTGATGATGATCTCGAGATCCAGT.....ACCG 2949
157 yfLeuValIleArgGluIleLeu..... 164
2950 CCCTCGGGCGCGGCTCATGTGACTTCTCTGAACAACCAACAGCTGCGCGCGG 2999
165GlnAlaSerAspLeuAlaGlyGlyCysArgValGluTh 177
3000 TGGTGAAGCGGTCTTATACCACTGCTCCACGAGAGAGAGAGGCCAGG 3049
177 rValAlaTrpSerSerAlaProGlyIleSerSerAsnLysSerSerThr 194
3050 AACGACGTGTCTTCCAGCCCA.....TCTCCGAGGAGAACGTGCGCGGA 3093
194 eu.SerSerThrHisGlnGlyLeuAlaTyrgIy..PheGluLeuSerTr 209
3094 TGTGACAGCCCTGAACGTGAGCAGCGCTGAAGGCTTACTTCAAGATGCGATG 3143

```
209 pLysArgAsnLeu.....Leu. 214
    |||
3144 GCTACAAGGGCTACGACCTGTCTACAGCCCCAGAGCGGCTTCACCTGC 3193
215 CysArgAsnCysAspArg.....SerArgGlyGlyGlyCysThrI 228
    |||
3194 GTGTCCCCGTCAGTAGGGGCTACTGTGACCATGAGGCCAGTGCACA 3243
228 eGlUglUAsnSerAspArgAlaThrCys.....ArgT 239
    |||
3244 CTGCCCCAGTGGGCGCCGTCGACGTGTGTCTCTCCATCTACACGG 3293
239 yRTpCysLysGluAspIleHisValSerLysLeuThrPheArgCysLys 255
    |||
3294 CTGGGGCGAGCAGCTGTGACACCTGAGC..... 3322
256 ValGluTYrTYrSerValTYrValLeuPhePheGlyGlyIleGly 272
    |||
3323 ATGAACACTGACGCGCTTCTCGCATCTTTGGG.....GCCCTGGG 3366
272 yGlyValLeuAla**ArgPheLeuLeuGlyIleProIleLeuIleAla 289
    |||
3367 CGGCCTCTTG.....CTGCTGGGGGTCGGGACGTTCTGTGTCC 3404
289 laValValTrpGlnCys 294
    |||
3405 TGGCGCTTCTGGGTTGC 3421
```

seq_name: /SIDSl/gcgdata/hold-geneseq/geneseqn-emb1/NA2001B.DAT:AA572520

seq_documentation_block:

ID AAS72520 standard; cDNA; 3875 BP.

XX AAS72520;

DT 13-FEB-2002 (first entry)

DE DNA encoding novel human diagnostic protein #8324.

KW Human; chromosome mapping; gene mapping; gene therapy; forensic;

KW food supplement; medical imaging; diagnostic; genetic disorder; ss.

XX Homo sapiens.

PN WO200175067-A2.

PD 11-OCT-2001.

PF 30-MAR-2001; 2001WO-US08631.

PR 31-MAR-2000; 2000US-0540217.

PR 23-AUG-2000; 2000US-0649167.

PA (HYSE-) HYSEQ INC.

PI Drmanac RT, Liu C, Tang YT;

DR MPI; 2001-639362/73.

DR P-PSDB; ABG08333.

PT New isolated polynucleotide and encoded polypeptides, useful in
PT diagnostics, forensics, gene mapping, identification of mutations
PT responsible for genetic disorders or other traits and to assess
PT biodiversity

PS Claim 1; SEQ ID No 8324; 103pp; English.

XX The invention relates to isolated polynucleotide (I) and
CC polypeptide (II) sequences. (I) is useful as hybridisation probes,
CC polymerase chain reaction (PCR) primers, oligomers, and for chromosome
CC and gene mapping, and in recombinant production of (II). The
CC polynucleotides are also used in diagnostics as expressed sequence tags

CC for identifying expressed genes. (I) is useful in gene therapy techniques
CC to restore normal activity of (II) or to treat disease states involving
CC (II). (II) is useful for generating antibodies against it, detecting or
CC quantitating a polypeptide in tissue, as molecular weight markers and as
CC a food supplement. (II) and its binding partners are useful in medical
CC imaging of sites expressing (II). (I) and (II) are useful for treating
CC disorders involving aberrant protein expression or biological activity.
CC The polypeptide and polynucleotide sequences have applications in
CC diagnostics, forensics, gene mapping, identification of mutations
CC responsible for genetic disorders or other traits to assess biodiversity
CC and to produce other types of data and products dependent on DNA and
CC amino acid sequences. AA564197-AA594564 represent novel human
CC diagnostic coding sequences of the invention.
CC Note: The sequence data for this patent did not appear in the printed
CC specification, but was obtained in electronic format directly from WIPO
CC at ftp.wipo.int/pub/published_pct_sequences.

XX SQ Sequence 3875 BP; 824 A; 1213 C; 1060 G; 778 T; 0 other;

alignment_scores:

Quality:	95.00	Length:	341
Ratio:	0.605	Gaps:	20
Percent Similarity:	46.041	Percent Identity:	20.821

alignment_block:

US-09-805-919-1 x AAS72520 ..

Align seg 1/1 to: AAS72520 from: 1 to: 3875

```
29 CysValPro.....SerSerCysGlyAspIleGlnIleLysPh 41
    |||
2496 TGGCTTCTGGGAAGGCGCTGCGAGCCCTGC..... 2525
41 ePropheArgLeuArgThrAspProGluHisCysGlyArgArgGlyTyrG 58
    |||
2526 .CCTCCAACCTGACTGGGGAATGGCGGACACTGCGGCTCTGGGAGCT 2574
58 luleAspCysGlnAsnAsnGlnThrValPheAsnTyr..... 70
    |||
2575 CTTCTCTGTGTGAGAACCAGTCTGCCCCGTGTAATTAAGTCTACATCAA 2624
70 ..... 70
2625 GGCCACTGCTACATCTCCAGACTCTGGGCTGTACGCCATGTGCACCTG 2674
71 .....LysSerArgIlePheAspValGlnIleAsnT 82
    |||
2675 CCCCCAGCCTTCACTGACAGCCGCTGCTC.....CTGGCTG 2712
82 yRArgSerTyrSer.....IleArgLeuLeu 90
    |||
2713 GGAACAACCTTCACTGCAACTGTCAACCTAGAACCTTCCCTTAAGAGTCAATC 2762
91 AspProGlyLeuAsnAspGlnArgGluAsnCysThrValPheProAsnH 107
    |||
2763 CAGCTCTGTCTCAGTGAAGAGAAATGCCCTCCATGGCAGAGTCAACGC 2812
107 sArgAlaSerTyrAspAlaMetThrSerGlnIlePheGluTrpValArgV 124
    |||
2813 CTGGGTGGCATACAGACTGGGAGACCCCTGGACATGGCGGCTTCTCCGC. 2861
124 alaAsnAsnAspIleAsnTyrValAsnCysLeuAlaProIleAsnSerSer 140
    |||
2862 ..AACAGCCAAAGTGAACGAATGATTTCTGCAGCACCGGCTCGGGAAGC 2909
141 GlnTYrIlePro.ThrSerPheCysSerLysAsnSerThrGlyPheSerT 157
    |||
2910 CCC..ATCCAACACTGATGATGATCTCTGGAGTTCCAGT.....ACCG 2950
157 yRLeuValIleArgGluIleLeu..... 164
    |||
2951 CCTCTGGGGGCCGCTCATGTGACTTCTCTGAACAACCAAGCTGTGCGCGG 3000
```

```

165 .....GlnAlaSerAspLeuAlaGlyCysArgValGluTh 177
      ::::: ||| |||
3001 TGGTGAGGCGCTTCTTATACCACGCTCCACGAGAGTAGAGGCCAGG 3050
      ::::: ||| ||| ||| |||
177 rValAlaTrpSerSerAlaProGlyIleSerSerAsnLysSerSerThrL 194
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3051 AACGACGTGGTCTTCCAGCCCA.....TCTCCGAGGAAGACGTGCGCGA 3094
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
194 eu.SerSerThrHisGlnGlyLeuAlaTyGly..PheGluLeuSerTr 209
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3095 TGTGACAGCCCTGAACGTGAGCAGCAGCTGAGGCTTACTTCAGATGCGATG 3144
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
209 pLysArgAsnLeu.....Leu. 214
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3145 GCTACAAGGGCTACGACCTGGTCTACAGCCCCAGAGCGGCTTCACCTGC 3194
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
215 CysArgAsnCysAspArg.....SerArgGlyGlyGlyCysThrI 228
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3195 GTGTCCCGGTGCAGTAGGGGCTACTGTGACCATGAGGCCAGTGCACAACA 3244
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
228 egluGluAsnSerAspArgAlaThrCys.....ArgT 239
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3245 CTTGCCCAGTGGGCCCCCTGCAGCTGTGTGTCCTTCTCCATCTACACGG 3294
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
239 yTrpCysLysGluAspIleHisValSerLysLeuThrPheArgCysLys 255
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3295 CTTGGGGCGAGCACTGTGACACCTGAGC..... 3323
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
256 ValGluTyrTyrSerValTyrValLeuPhePheGlyGlyIleGlyIleG 272
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3324 ATGAAACTCGACGCGTCTCTCGCATCTCTTTGGG.....GCCCTGGG 3367
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
272 yGlyValLeuAla**ArgPheLeuLeuGlyIleProIleLeuIleAla 289
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3368 CGGCCTCTG.....CTGCTGGGGGTCGGGACGTTGCTGTCTCC 3405
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
289 laValValTrpGlnCys 294
      ::::: ||| ||| ||| ||| ||| ||| ||| ||| |||
3406 TGGCCTTCTGGGGTGC 3422

```



```
22 sAlaGlnAspIleSerGlnCysValProSerSerCysGlyAspIleGlnI 39
| :::::::::::::: :::: ::::| ::::| ::::|
4149 A...AATGACATTAAGCAAACTCATCTTCAAAAGT...GGAGATTTCCAA 4192
39 le.....LysPheProPheArgLeuArgThrAspProGluHis 51
::: ::::| ::::| ::::| ::::|
4193 TGTCCTCATATGCTGAATACCCCAAAATCTCGCTCA.....CAAGAA 4236
52 CysGlyArgArgGlyTyrGluLeuAspCysGlnAsnaenGlnThrValph 68
| ::::| ::::| ::::| ::::|
4237 TGGGGCCGAGAGCTATGAAATGCAT...GAGAATGGAAGTACC..... 4278
68 eAsnTyrIlySerArgIlePheAspValGlnGluIleAsnTyrArgSert 85
::: ::::| ::::| ::::| ::::|
4279 .....AAAACTCTCTCCAGATGACGGATGTGTACTACTCGCCTA 4318
85 yrSerIleArgLeuLeuAspProGlyLeuAsnaSpGlnArgGluAsnCys 101
| ::::| ::::| ::::| ::::|
4319 CAAGTGTAGG.....AATCCAGAACTTGAACGAAC 4350
102 ThrValPheProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnI 118
::: ::::| ::::| ::::| ::::|
4351 GGACTCTACCCGCTTACACTGGACTGCCAGATCACGGCATCTTGCAT 4400
118 ePheGluTyrValArgValAsnaSpIleAsnTyrValAsnCysLeuA 135
| ::::| ::::| ::::| ::::|
4401 TTTC..... 4404
135 laProIleAsnSerSerGlnTyrIleProThrSerPheCysSerLysAsn 151
| ::::| ::::| ::::| ::::|
4405 .....CCCGACAGTATACCCG...TCTTCATCAGTGATGAA 4440
152 SerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAlaSerAs 168
| ::::| ::::| ::::| ::::|
4441 AGCAGAGAAGAGACTACTTTTAAGTCCAGAG..... 4473
168 pleuAlaGlyGlyCysArgValGluThrValAlaTyrPheSerAlaProg 185
| ::::| ::::| ::::| ::::|
4474 .....AGAGAGGACTCATTTGCTGTGAGCCAGTCACCTG 4507
185 lyIleSerSerAsnLysSer..... 191
| ::::| ::::| ::::| ::::|
4508 GGACCTCTGTCTCAGAGGACCGCACCAGAGCGCTGCCGCCAGATTGTGC 4557
192 ..SerThrLeuSer...SerThrHisGlnGlyLeuAlaTyrGlyPheG 206
| ::::| ::::| ::::| ::::|
4558 GGAGCCACGCTGAGTGGCAAGCAGAGAGGACAGGACATGCGGGCGTG 4607
206 uLeuSerTyrPlySarg.....AsnL 213
| ::::| ::::| ::::| ::::|
4608 ACCACAGTGGAGAGAGAGAGTGGATGGAAACACACAGGCTGCTCATTCAG 4657
213 eu.LeuCysArgAsnCysAspArgSerArgGly..... 223
| ::::| ::::| ::::| ::::|
4658 CACCTTTGTTGTTACTGTGAACGTGATGTGGGCCAGTATCAAGAGAGTC 4707
224 .....G 224
4708 TCTCTGAGTACTGCACCATGCACTGCACACGAGGCGACTATTAGCCAG 4757
224 lyGluCysThrIleGluGluAsnSerAspArgAlaThrCysArg.TyrTr 240
| ::::| ::::| ::::| ::::|
4758 GGCAGAGACCACT.....AGACTTCAGTGCAGGGAGACCTG 4789
240 pCysIlySgluAspIleHisValSerLysLeuThrPheArgCysIlyValG 257
| ::::| ::::| ::::| ::::|
4790 GTTTCCCTCGTTTGCACCTTAGTAATAATTGGGTGGAGG.....TTTC 4833
257 luIyrrTyrSerValTyrValLeuphe 265
| ::::| ::::| ::::| ::::|
4834 CTITTGATCTGTGTTTGAAGACTGTTC 4859
seq_name: /cgn2_6/ptodata/2/ina/5B_COMB.seq:US-08-599-654-7
```

```
seq_documentation_block:
; Sequence 7, Application US/08599654
; Patent No. 5882925
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITL OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/599,654
; FILING DATE: 09-FEB-1996
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; APPLICATION NUMBER DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-041
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6407 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; US-08-599-654-7

alignment_scores:
Quality: 94.50 Length: 310
Ratio: 0.690 Gaps: 18
Percent Similarity: 44.194 Percent Identity: 23.226

alignment_block:
US-09-805-919-1 x US-08-599-654-7 ..

Align seg 1/1 to: US-08-599-654-7 from: 1 to: 6407

6 LeuLeupheAlaCysValLeuLeuValThrLeuIleSerSerLys 22
| ::::| ::::| ::::| ::::|
4099 CTGCTCATCTAGGCAATGCACCTGATTGTTACTCTGTTCAGAAAGATAA 4148
22 sAlaGlnAspIleSerGlnCysValProSerSerCysGlyAspIleGlnI 39
| ::::| ::::| ::::| ::::|
4149 A...AATGACATTAAGCAAACTCATCTTCAAAAGT...GGAGATTTCCAA 4192
39 le.....LysPheProPheArgLeuArgThrAspProGluHis 51
| ::::| ::::| ::::| ::::|
52 CysGlyArgArgGlyTyrGluLeuAspCysGlnAsnaenGlnThrValph 68
| ::::| ::::| ::::| ::::|
4193 TGTCCTCATATGCTGAATACCCCAAAATCTCGCTCA.....CAAGAA 4236
```

```
4237 TGGGGCCGAGAGCTATTGAATGCAT...GAGATGGAAGTACC..... 4278
68 easnTyrlYsSerArgIlePheAspValGlnGluIleasnTyArgSerT 85
4279 .....AAAACCTCCTCCAGATGACGATGTGTACTACTCGCCTA 4318
85 yRSerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCys 101
4319 CAAGTGTAGG.....AATCCAGAACTTGAACGAAAC 4350
102 ThrValPheProAsnHisArgAlaSerTyArgPalamethrSerGlnI 118
4351 GGACTCTACCCGGCCTACACTGAGTCCAGGATCACGGCATCTTGCAT 4400
118 ePheGluTrpValArgValAsnAspIleasnTyArgValAsnCysLeu 135
4401 TTTC..... 4404
135 IAProIleAsnSerSerGlnTyrlleProThrSerPheCysSerLysAsn 151
4405 .....CCCGACAGTATAACCCG...TCCTTCATCAGTATGAA 4440
152 SerThrGlyPheSerTyrlleuValIleArgGluIleLeuGlnAlaSer 168
4441 AGCAGAAGAGAGACTATTTAAGTCCAGAG..... 4473
168 PleuAlaGlyGlyCysArgValGluThrValAlaTrpSerSerAlaPro 185
4474 .....AGAGAGGAGACTATGCTCTGAGCCAGTACACTG 4507
185 lyIleSerSerAsnLysSer..... 191
4508 GGACCTCTGCTCAGAGACCGCACACAGAGGCTGCCAGATTGTCG 4557
192 ...SerThrLeuSer...SerThrHisGlnGlyLeuAlaTyArgIlePhe 206
4558 GGAGCCACGCTGAGTGCACAAGAGAGGACAGGCATGCGGGCGCTG 4607
206 uLeuSerTrpLysArg.....AsnL 213
4608 ACCACAGTGGAGAGACAGCTGATGTGAACACAGGCTGCTCATTCAG 4657
213 euLeuCysArgAsnCysAspArgSerArgIly..... 223
4658 CACCTTGTGTGTACTGTGAACGTGAATGTGGGCCAGTATCAAGAGTTC 4707
224 .....G 224
4708 TCTGTAGTGACTGCACCATGCACTGCACACGAGGCGACTATTAGCCAG 4757
224 lyGluCysThrIleGlnGluAsnSerAspArgAlaThrCysArg.TyrTr 240
4758 GGCAGACCACCT.....AGACTTCAGTGCAGGAGCCTG 4789
240 pCysLysGluAspIleHisValSerLysLeuThrPheArgCysLysVal 257
4790 GTTTCCTCCTCGTTTGCACCTTTAGTAATTTGGGTGGCAGG.....TTTC 4833
257 LuTyrlTyrservalTyArgValLeuPhe 265
4834 CTTTGGATCTGTTTGAAGCTGTTT 4859

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-944-868A-7

seq_documentation_block:
; Sequence 7, Application US/08944868A
; Patent No. 6018025
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
```

```
ADDRESSEE: PENNIE & EDMONDS
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/944,868A
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/599,654
FILING DATE:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/386,844
FILING DATE: 10-FEB-1995
ATTORNEY/AGENT INFORMATION:
NAME: CORUZZI, LAURA A
REGISTRATION NUMBER: 30,742
REFERENCE/DOCKET NUMBER: 7853-041
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 6407 base pairs
TYPE: nucleic acid
STRANDEDNESS: both
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
US-08-944-868A-7

alignment_scores:
Quality: 94.50 Length: 310
Ratio: 0.690 Gaps: 18
Percent Similarity: 44.194 Percent Identity: 23.226

alignment_block:
US-09-805-919-1 x US-08-944-868A-7 ..

Align seg 1/1 to: US-08-944-868A-7 from: 1 to: 6407

6 LeuLeuPheAlaCysValLeuLeuValThrLeuIleSerSerly 22
|||||: : : : : |||: : : : :
4099 CTGCTCATCTGAGGATCGACCTGATGTTACTGTTCGAGAAAGATAA 4148
22 sAlaGlnAspIleSerGlnCysValProSerSerCysGlyAspIleGln 39
| : : : : |||: : : : |||: : : :
4149 A...AATGACATTAAGCAAACTCATCTTCAAAAGT...GGAGATTTCCAA 4192
39 le.....LysPheProPheArgLeuArgThrAspProGluHis 51
: : : : : |||: : : : :
4193 TGTCCCATATGCTGAATACCCCAAAATCCTCGCTCA.....CAAGAA 4236
52 CysGlyArgArgGlyTyrlleuLeuAspCysGlnAsnAsnGlnThrVal 68
|||||: : : : : |||: : : : :
4237 TGGGGCCGAGAGCTATTGAATGCAT...GAGATGGAAGTACC..... 4278
68 easnTyrlYsSerArgIlePheAspValGlnGluIleasnTyArgSerT 85
4279 .....AAAACCTCCTCCAGATGACGATGTGTACTACTCGCCTA 4318
85 yRSerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCys 101
|||||: : : : : |||: : : : :
4319 CAAGTGTAGG.....AATCCAGAACTTGAACGAAAC 4350
```



```

135 laProtleasnserserGlnTyrIleProThrSerPheCysSerLysAsn 151
      ::::::::::: ||| ||| ||| ::::
4405 .....CCCGACAGTATAACCG...TCTTCATCAGTGATGAA 4440
152 SerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAlaSerAs 168
      ||| ::::::::::: ||| ||| |||
4441 AGCAGAGAGAGAGACTACTTTTAAGTCCAGAG..... 4473
168 pleuAlaGlyGlyCysArgValGluThrValAlaTrpSerSerAlaProG 185
      ||| ||| ::::::::::: ||| ||| ||| |||
4474 .....AGAGAGGAGTCAATTGCTCTGAGCCAGTCACCTG 4507
185 lyIleSerSerAsnLysSer..... 191
      ||| ||| ::::::::::: ||| ||| ||| |||
4508 GGACCTCTGCTCAGAGAGACCGCACAGAGGCTGCGCCAGATTGTCTG 4557
192 ...SerThrLeuser...SerThrIleGlnGlyLeuAlaTyrGlyPheG1 206
      ::::::::::: ||| ||| ||| ||| |||
4558 GGAGCCACGCTGAGTGCAAGCAGAGAGAGGACAGGCAATGCGGGCGTG 4607
206 uLeuSerTrpLysArg.....AsnL 213
      ::::::::::: ||| ||| ||| ||| |||
4608 ACCACAGTGGAGAGACAGGTGATGTGAACACACAGGCTGCTCATTCAG 4657
213 eu.LeuCysArgAsnCysAspArgSerArgLy..... 223
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4658 CACCTTTGTTGTTACTGTGAACGTGAATGTGGCCACAGTATCAAGAGAGTC 4707
224 .....G 224
4708 TCTCTGAGTGAAGTGCACCATGGCAGTGGCAGGCGACTATTAGCCAG 4757
224 LyGluCysThrIleGluGluAsnSerAspArgAlaThrCysArg.TyrTr 240
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4758 GGCAGAGCACT.....AGACTTCAGTGCAGGAGCACTG 4789
240 pCysLysGluAspIleHisValSerLysLeuThrPheArgCysLysValG 257
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4790 GTTTCCCTCGTTTGCACTTAGTAATTTGGGTGGAGG.....TTTC 4833
257 LuTyrTrpSerValTyrValLeuPhe 265
      ::| ||| ||| ||| ||| ||| ||| ||| |||
4834 CTTTGGATCTGTTTGAGACTGTTC 4859

seq_name: /cgn2_6/ptodata/2/ina/6A_COMB.seq:US-08-944-496-7

seq_documentation_block:
; Sequence 7, Application US/08944496
; Patent No. 6124433
; GENERAL INFORMATION:
; APPLICANT: FALB, DEAN A
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE
; TITLE OF INVENTION: TREATMENT AND DIAGNOSIS OF CARDIOVASCULAR DISEASE
; NUMBER OF SEQUENCES: 54
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PENNIE & EDMONDS LLP
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/944,496
; FILING DATE: 06-OCT-1997
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/599,654

```

```

; FILING DATE: 09-FEB-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/485,573
; FILING DATE: 07-JUN-1995
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/386,844
; FILING DATE: 10-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: CORUZZI, LAURA A
; REGISTRATION NUMBER: 30,742
; REFERENCE/DOCKET NUMBER: 7853-104
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6407 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: both
; TOPOLOGY: unknown
; MOLECULE TYPE: CDNA
; US-08-944-496-7

alignment_scores:
      quality: 94.50      length: 310
      ratio: 0.690      gaps: 18
      percent similarity: 44.194      percent identity: 23.226

alignment_block:
US-09-805-919-1 x US-08-944-496-7 ..

Align seg 1/1 to: US-08-944-496-7 from: 1 to: 6407

6 LeuLeuPheAlaCysValLeuLeuValThrLeuIleSerSerSerLy 22
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4099 CTGCTATCTTAGGCATCGCACTGATTGTTACCTGTTCAGAAAGATTA 4148
22 sAlaGlnAspIleSerGlnCysValProSerSerCysGlyAspIleGlnI 39
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4149 A..AATGACATAAGCAAACTCATCTCAAAAGT...GGAGATTCCAAA 4192
39 le.....LysPheProPheArgLeuArgThrAspProGlnHis 51
      ::| ||| ||| ||| ||| ||| ||| ||| |||
4193 TGTCCCATATGCTGAATATCCCAAAATCTCGCTCA....CAAGAA 4236
52 CysGlyArgArgGlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPh 68
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4237 TGGGGCCGAGAAAGCTATTGAATGCAT..GAGAAATGGAAGTACC.... 4278
68 eAsnTyrLysSerArgIlePheAspValGlnGluIleAsnTyrArgSerT 85
      ::| ||| ||| ||| ||| ||| ||| ||| |||
4279 .....AAAACTCTCTCCAGATGACGAGATGTACTACTCGCCTTA 4318
85 yrSerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCys 101
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4319 CAAGTGTAAG.....AATCAGAACTTGAACGAAC 4350
102 ThrValPheProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnI 118
      ::| ||| ||| ||| ||| ||| ||| ||| |||
4351 GGACTCTACCCGGCTTACACTGACTGCCAGAGATCAGGCAATCTTGCA 4400
118 epheGluTrpValArgValAsnAsnAspIleAsnTyrValAlaAsnCys 135
      ||| ||| ||| ||| ||| ||| ||| ||| |||
4401 TTTC..... 4404
135 laProtleasnserserGlnTyrIleProThrSerPheCysSerLysAsn 151
      ::| ||| ||| ||| ||| ||| ||| ||| |||
4405 .....CCCGACAGTATAACCG...TCTTCATCAGTGATGAA 4440
152 SerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAlaSerAs 168
      ||| ||| ||| ||| ||| ||| ||| ||| |||

```



```
4441 AGCAGAGAGAGACTACTTTTAAGTCCAGAG..... 4473
168 pLeuAlaGlyGlyCysArgValGluThrValAlaTrpSerSerAlaProG 185
4474 .....AGAGAGGAGACTCATCTGCTGAGCCAGTCACTG 4507
185 lYleSerSerAsnLysSer..... 191
4508 GGACCTCTGCTCAGAGAGACCGCAGCAGAGCGCTGCCGCCAGATTTGTCG 4557
192 ...SerThrLeuSer...SerThrHisGlnGlyLeuAlaTyrglyPheG 206
4558 GGAGCCACGCTGAGTGGCAAGCAGAGAGGAGACAGCATGCGGGCGTG 4607
206 uLeuSerTrpLysArg.....AsnL 213
4608 ACCACAGTGGAGAGAGACAGGTGATGTGGAACACAGCGCTCATTTGAG 4657
213 eu.LeuCysArgAsnCysAspArgSerArgGly..... 223
4658 CACCTTTGTTGTTACTGTGAACGTGAATGTGGGCCAGTATCAAGAGAGTC 4707
224 .....G 224
4708 TCTCTGAGTACTGCACCATGCGACTGGCACAGGGCGACTATTAGCCAG 4757
224 lYgluCysThrIleGluGluAsnSerAspArgAlaThrCysArg.TyrTr 240
4758 GGCAGACCACT.....AGACTTCAGTGCAGGAGCCTG 4789
240 pCysLysGluAspIleHisValSerLysLeuThrPheArgCysLysValG 257
4790 GTTTCCCTTCGTTTTCACCTTGAATTAATTGGGTGGAGG.....TTTC 4833
257 lUtyrTySerValTyrValLeuPhe 265
4834 CTTTGGATCTTTTGGAGACTGTTC 4859
seq_name: /cgn2_6/ptodata/2/ina/6a_COMB.seq:US-08-434-000A-3
seq_documentation_block:
; Sequence 3, Application US/08434000A
; Patent No. 6046037
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; APPLICANT: K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; TITLE OF INVENTION: PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Lyon & Lyon
; STREET: 633 West Fifth Street
; STREET: Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; MEDIUM TYPE: storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/434,000A
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA: including application
; PRIOR APPLICATION DATA: described below: 1
; APPLICATION NUMBER: 08/367,395
; FILING DATE: 12/30/94
; ATTORNEY/AGENT INFORMATION:
```

```
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
; TELEX: 67-3510
; TELEX: SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2919 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 235.....2472
; US-08-434-000A-3

alignment_scores:
Quality: 92.50 Length: 304
Ratio: 0.642 Gaps: 22
Percent Similarity: 47.368 Percent Identity: 25.658

alignment_block:
US-09-805-919-1 x US-08-434-000A-3 ..
Align seg 1/1 to: US-08-434-000A-3 from: 1 to: 2919
29 CysValProSer.....SerCysGlyAspIleGlnIleLysPheProPh 43
82 TGTGTCCCTGAAGTCAGATCAACGGAGAGAGAAAGTGGCTAAACATT 131
43 eArgLeuArgThrAspProGluHisCysGlyArgArgGly..... 56
132 GCACAGAGAGAGTGGCGCTGAGTGTGGCGGCTCGGAGACCACAGCAA 181
57 .....TyrGluLeuAspCysGlnAsnAsnGlnThrValPhe..... 68
182 TGCTGCTCTTCTGCTGCTACCTGCTG.....CTGGCGTCTTCCAGCC 225
69 ...AsnTyrLysSerArgIlePheAspValGlnIleAsnTyr..... 82
226 ATCTCCACGAAGAGTCCCATATTGTTGTCGCCAGAGAGTGAATAGTGTGA 275
83 ...ArgSerTyrSerIleArgLeuLeuAspProGlyLeuAsnAspGlnA 98
276 AGGTAACTAGTGTCCATCAGTGTCTACTACCAACCACCTCTGTCAAC 325
98 rgGluAsnCysThrValPheProAsnHisArgAlaSerTyrAspAlaMet 114
326 GG.....CACACCCGGAAGTAC..... 342
115 ThrSerGlnIlePheGluTrpValArgValAsnAsnAspIleAsnTyrVa 131
343 .....TGSTGCCGCGCAGGGA.....GCTAGAGG 365
131 lAsnCysLeuAlaProIleAsnSerSerGlnTyrIleProThrSerPheC 148
366 TGGCTGATACCTCATCTCTCGGAGGGCTACGTCTCCAGCAATATATG 415
148 ysSerLys..AsnSerThrGlyPhe.....SerTyrLeuVal 159
416 CAGGACAGGGCTAACCTCACCACACTTCCCGAGAGAGGCACATTTGTGTG 465
160 lLeArgGluIleLeuGlnAlaSerAsp.LeuAlaGlyGlyCysArgValG 176
466 .....AACATTGCCAGCTGAGCCAGAGATGACTCCGGGC.....GCTA 503
176 lUthrValAlaTrpSerSer.....AlaProGlyLysSerSer 188
```

```
504 CAAGTGTGGCTGGGCATCAATAGCCGAGGCTGTCTTTGATGTACACC 553
189 AsnlysserSer.....ThrleuSerSerThrHl 198
      :::::
554 TGGAGGTACGCCAGGGTCTGGGCTCTAAATGACACTAAAGTCTACA.. 601
198 sGlnGlyleuAlaTyrGlyPheGlnleuSerTrpLysArg..... 211
      :::::
602 .....CAGTGGACCTGGGCAGAACGGGTGACCAT 629
212 ..AsnleuLeuCysArgAsnCysAspArgSerArgGlyGlyCysThr 227
      :::::
630 CAAGTGTCCCTTTCAGAGACTGAGAATGCTCAAAAGAGAGAGTCTGTACA 679
228 TleGlnGluAsnSerAspArgAlaThrCysArgTyr..TrpCysLysGl 243
      :::::
680 .....AGCAGATAGGCCCTGTACCCCTGTGCTGTGTCATCGAC 714
243 uAspIle.....HisValSerLysLeuThrP 252
      :::::
715 TCCAGTGGTTATGTGAATCCCACTATACAGAGAAATACGCCCTTGATAT 764
252 heArgCysLysValGluTyrTyrSerValTyrValleuPheGly 268
      :::::
765 TCAGGGTACTGGCCAGTACTGTTCAGCGTGTGCATCAACCACTCAGGC 814
269 TleGlyTleGlyGlyValleuAla**ArgPheLeuLeuGly..IlePr 284
      :::::
815 TCAGCGATGCTGGGAGATCTCTGCCAGG.....CTGGGATGATGCC 858
284 OileuIle 287
      :::::
859 AATAGTAATA 868
```

seq_name: /cgn2_6/ptodata/2/lna/6B_COMB.seq:US-09-312-157-3

```
seq_documentation_block:
; Sequence 3, Application US/09312157
; Patent No. 6303341
; GENERAL INFORMATION:
; APPLICANT: ANDREW C. HIATT, JULIAN
; K.-C. MA, THOMAS LEHNER
; TITLE OF INVENTION: IMMUNOGLOBULINS CONTAINING PROTECTION
; PROTEINS IN PLANTS AND THEIR USES
; NUMBER OF SEQUENCES: 19
; CORRESPONDENCE ADDRESS:
; ADDRESS: Lyon 6 Lyon
; STREET: 633 West Fifth Street
; Suite 4700
; CITY: Los Angeles
; STATE: California
; COUNTRY: U.S.A.
; ZIP: 90071
; COMPUTER READABLE FORM:
; MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
; storage
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: IBM P.C. DOS 5.0
; SOFTWARE: Word Perfect 5.1
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/312,157
; FILING DATE: 14-May-1999
; CLASSIFICATION: <unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/434,000
; FILING DATE: <unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Guise, Jeffrey W.
; REGISTRATION NUMBER: 34,613
; REFERENCE/DOCKET NUMBER: 212/127
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (619) 552-8400
; TELEFAX: (619) 552-0159
```

```
;
; TELEX: 67-351
; SEQUENCE LISTING
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2919 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 235...2472
; SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-312-157-3
```

alignment_scores:

Quality:	92.50	Length:	304
Ratio:	0.642	Gaps:	22
Percent Similarity:	47.368	Percent Identity:	25.658

alignment_block:

US-09-805-919-1 x US-09-312-157-3 ..

Align seg 1/1 to: US-09-312-157-3 from: 1 to: 2919

```
29 CysValProSer.....SerCysGlyAspIleGlnIleLysPheProph 43
      :::::
82 TGTGTCCCTGAAGTCAGATCAACGGGAGAGAAAGAGTGGCTAAACATT 131
43 eaArgLeuArgThrAspProGluHisCysGlyArgGly..... 56
      :::::
132 GCACAGAGAGAGAGTGGCTGAGTGTGGCGGCGCTCGGACCCACAGCAA 181
57 .....TyrGluLeuAspCysGlnAsnAsnGlnThrValPhe..... 68
      :::::
182 TGCTGCTCTTCTGCTCACCCTGCTG.....CTGGCGGTCTTCCACAGCC 225
69 ..AsnTyrLysSerArgIlePheAspValGlnIleAsnTyr..... 82
      :::::
226 ATCTCCACGAAGAGTCCCATATTGTGCCGAGAGAGTGAATAGTGTGA 275
83 ...ArgSerTyrSerIleArgLeuLeuAspProGlyLeuAsnAspGlnA 98
      :::::
276 AGGTAAGTCAAGTGTCCATCAGCTGCTACTACCAACCCACCTCTGTCAAC 325
98 rgGluAsnCysThrValPheProAsnHisArgAlaSerTyrAspAlaMet 114
      :::::
326 GG.....CACACCCGGAGATAC..... 342
115 ThrSerGlnIlePheGluTrpValArgValAsnAsnAspIleAsnTyrVa 131
      :::::
343 .....TGTCGCCGGCAGGA.....GCTAGAG 365
131 IAsnCysLeuAlaProIleAsnSerSerGlnTyrIleProThrSerPheC 148
      :::::
366 TGGCTGCATAACCCCTCATCTCTCGAGGGCTACGCTCCAGCAATATG 415
148 ysSerLys..AsnSerThrGlyPhe.....SerTyrIleuVal 159
      :::::
416 CAGGCAGGGCTAACCTCAACCAACTTCCCGGAGAACGGCACATTTGTGTG 465
160 IleArgGluIleLeuGlnAlaSerAsp..LeuAlaGlyGlyCysArgValG 176
      :::::
466 .....AACATTGCCAGCTGAGCCAGATGACTCCGGG.....GCTA 503
176 luThrValAlaTrpSerSer.....AlaProGlyIleSerSer 188
      :::::
504 CAAGTGTGGCTGGGCATCAATAGCCGAGGCTGTCTTTGATGTACAGCC 553
189 AsnlysserSer.....ThrleuSerSerThrHl 198
      :::::
554 TGGAGGTACGCCAGGGTCTGGGCTCTAAATGACACTAAAGTCTACA.. 601
```

```
198 sGInGlyLeuAlaTyRgLyPheGluLeuSerTrpLysArg..... 211
      :::: :::: ::::
602 .....CAGTGGACCTGGGAGAGCGGTGACCAT 629
212 ..AsnLeuLeuCysArgAsnCysAspArgSerArgGlyGlyGluCysThr 227
      :::: :::: ::::
630 CAACTGCCCTTCAGAGACTGAGAAATGCTCAAAAAGAGAGAGTCTGTACA 679
228 IleGluGluAsnSerAspArgAlaThrCysArgTyr..TrpCysLysG1 243
      :::: :::: ::::
680 .....AGCAGATAGGCTGTACCCCTGTGCTGTGCATCGAC 714
243 uAspIle.....HisValSerLysLeuThrP 252
      :::: ::::
715 TCCAGTGGTTANGTAATCCCACTATACAGGAAGATACGCCCTTGATAT 764
252 heArgCysLysValGluTyrTyrSerValTyrValLeuPhePheGlyGly 268
      :::: ::::
765 TCAGGGTACTGGCCAGTACTGTTTCAGCGTTGTCATCAACCACTCAGGC 814
269 IleGlyIleGlyValLeuAla**ArgPheLeuLeuGly...IlePr 284
      :::: ::::
815 TCAGCGATGCTGGCAGTATCTCTGCCAGG.....CTGGGATGATTCC 858
284 oIleLeuIle 287
      :::: ::::
859 AATAGTATA 868
```

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-164-839-7

seq_documentation_block:

```
; Sequence 7, Application US/08164839
; Patent No. 5514573
; GENERAL INFORMATION:
; APPLICANT: YASUEDA, HISASHI
; APPLICANT: NAKANISHI, KAZUO
; APPLICANT: MOTOKI, MASAO
; APPLICANT: NAGASE, KAZUO
; APPLICANT: MATSUI, HIROSHI
; TITLE OF INVENTION: GENE ENCODING TRANSGlutAMINASE DERIVED
; TITLE OF INVENTION: FROM FISH
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,
; ADDRESSEE: P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/164,839
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/004,729
; FILING DATE: 14-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5514573man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-599-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)412-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
```

```
; LENGTH: 2088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Theragra chalcogramma
; TISSUE TYPE: liver
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2085
; US-08-164-839-7
```

alignment_scores:

Quality:	91.50	Length:	250
Ratio:	0.855	Gaps:	10
Percent Similarity:	42.800	Percent Identity:	21.600

alignment_block:

US-09-805-919-1 x US-08-164-839-7/rev ..

Align seg 1/1 to reverse of: US-08-164-839-7 from: 1 to: 2088

```
86 SerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThr 102
      :::: ::::
902 TCTGTGTCATCGGGCTGACTCAAGTTGTTGATGACGCGTGTGGG..... 858
102 rValPheProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnIleP 119
      :::: ::::
857 ...ATGCCAGGATCGCAGCACTGTGCAGGCCACGGCTGCAAAACACC 812
119 heGluTrpValArgValAsnAsp..... 127
      :::: ::::
811 AGCATTGGCCATCTTCACTCCAGACTTTGATTGCTCCACTTCACAGG 762
128 .....IleAsnTyrValAsnCysLeuAlaProIleAsnSe 139
      :::: ::::
761 ATGGCACACTGCTCATCCATCTTGGTTGGCTGACTCCACCTGTGAAG 712
139 rSerGlnTyrIleProThrPheCysSerLysAsnSerThrGlyPheS 156
      :::: ::::
711 TTCCTGCCATTCCCCACCAAG...ACTCCACCGTCATCGTTGCCATTCA 665
156 erTyrLeuValIle.....Arg 161
      :::: ::::
664 CCATGGCAACGACTGCCCTGCTGATGTAGACAGAGGTCCCATCGCTGACGC 615
162 GluIleGluAlaSerAspLeuAlaGly.....G1 172
      :::: ::::
614 ATGTGTTTGGGGGTCTGACTTGTGCTGCTTAAGCGGTCAAGATCTC 565
172 yCysArgValGluThrValAlaTyrSerSerAlaProGlyIleSerSerA 189
      :::: ::::
564 AAAGCAGATGTCCATCACATGTCTTCAAACTGTCCGAAATTCCATGGCA 515
189 snLys..... 190
      ::::
514 ACCCATCATCCATCCGCACTCCCACTGAAAAATGGAGACCATTCCTGCTTC 465
191 .....SerSer.....ThrLeuSerSe 196
      ::::
464 ATAATGTAATTCCTTGAGCAGACTTTCATCAGGAGGTACACAGCATCATC 415
196 rThrHisGlnGlyLeuAlaTyRgLyPheGluLeuSerTrpLysArgAsnL 213
      :::: ::::
414 TTGCACACAGGATTAACAGAGATGATTTTG.....GTAAGTCCG 371
213 euLeuCysArgAsnCys..... 218
      ::::
370 TCCTCTCCAGAGATTATTCCTGACATCAGTTCACAGCCAGCGGTAC 321
219 .....AsparGSerArgGlyGlyGluCysThrIleGluGluAsnSerAs 233
```

```

320 TCGCCAAGCGCAGCAGCGCAGCTGTACACAGTCAGCAGTAATTCGTT 271
233 pargalaThrCysArgTyrTrpCysLysGluAspIleHisValSerLysL 250
270 CCTCGCGCTCTGC.....TGCTGTAACCACTGTCACCAAGCCTC 227
250 eutHrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePhe 266
226 CAGCCGTCGCTGTCTGATCAACAATATCATCGTGTCTTATCCTGCTT 177
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:us-08-583-799-7
```

seq_documentation_block:

```

; Sequence 7, Application US/08583799
; Patent No. 5607849
; GENERAL INFORMATION:
; APPLICANT: YASUEDA, HISASHI
; APPLICANT: NAKANISHI, KAZUO
; APPLICANT: MOTOKI, MASAO
; APPLICANT: NAGASE, KAZUO
; APPLICANT: MATSUI, HIROSHI
; TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
; FROM FISH
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/004,729
; FILING DATE: 14-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5607849man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-599-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)412-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 7:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2088 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Theragra chalcogramma
; TISSUE TYPE: liver
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2085
; US-08-583-799-7
```

```

alignment_scores:
  Quality: 91.50      Length: 250
  Ratio: 0.855      Gaps: 10
  Percent Similarity: 42.800      Percent Identity: 21.600
```

alignment_block:

US-09-805-919-1 x US-08-583-799-7/rev ..

Align seg 1/1 to reverse of: US-08-583-799-7 from: 1 to: 2088

```

86 SerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysTh 102
902 TCTGTGTCATGGGCTGACTCAAAAGTTGGTATGATGACAGCGTGGGG... 858
102 rValPheProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnIleP 119
857 ...ATGCCAGGCATCGCAGCAGCTGTCCAGGCCACGGCTGCCAAACACCC 812
119 heGluTrpValArgValAsnAsp..... 127
811 AGCATTGGCCATCTACTTCACCTCCAGACTTTGATTGCTCCACTTCTCCAGG 762
128 .....IleAsnTyrValAsnCysLeuAlaProIleAsnSe 139
761 ATGGGCACACTGCTCATCTCATTTGGTGGCTGTACTCCACCTGTGTAAGG 712
139 rSerGlnTyrIleProThrSerPheCysSerLysAsnSerThrGlyPheS 156
711 TTCCTGCCATTGCCCAACCAAG..ACTCCACCGTCATCGTTGGCATTCA 665
156 erTyrLeuValIle.....Arg 161
664 CCATGGCAACGACTGCCCTGCTGATGTAGACAGGTCGCCATGCTGACGC 615
162 GluIleLeuGlnAlaSerAspLeuAlaGly.....G1 172
614 ATGTCGTTTGGGGGCTGACCTTGTCTGCTTAAAGCGGTCAAGATCTC 565
172 yCysArgValGluThrValAlaTrpSerSerAlaProGlyIleSerSera 189
564 AAAGCAGATGTCCATCAACATTTGTCTCAAACTGTCCGAATTCATGCA 515
189 snLys..... 190
514 ACCCACTCATCCAAATCCGCACTCCCAAGTGAATGCGACCATTCCTGTC 465
191 .....SerSer.....ThrLeuSerSe 196
464 ATAATGATATTCCTTGACGACAGCTTTCATCAGGGAGGTACACAGCATC 415
196 rThrHisGlnGlyLeuAlaTyrGlyPheGluLeuSerTrpLysArgAsnL 213
414 TTTCACACGAGGATTAACAGCAAGTACATTG.....GTAAAGTCCG 371
213 euLeuCysArgAsnCys..... 218
370 TCCTCTCCAGAGTTTATTCCTTGACATCATCTCAACAGCCAAGCGGTAC 321
219 .....AsparSerArgGlyGlyGluCysThrIleGluGluAsnSeraS 233
320 TCGCCAAGCGCAGCAGCGCAGCTGTACACAGTCAGCAGTAATTCGTT 271
233 pargalaThrCysArgTyrTrpCysLysGluAspIleHisValSerLysL 250
270 CCTCGCGCTCTGC.....TGCTGTAACCACTGTCACCAAGCCTC 227
250 eutHrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePhe 266
226 CAGCCGTCGCTGTCTGATCAACAATATCATCGTGTCTTATCCTGCTT 177
seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-164-839-9
```

```

seq_documentation_block:
; Sequence 9, Application US/08164839
; Patent No. 5514573
; GENERAL INFORMATION:
; APPLICANT: YASUEDA, HISASHI
```

```

: APPLICANT: NAKANISHI, KAZUO
: APPLICANT: MOTOKI, MASAO
: APPLICANT: NAGASE, KAZUO
: APPLICANT: MATSUI, HIROSHI
: TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
: NUMBER OF SEQUENCES: 72
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
: OPERATING SYSTEM: PC-DOS/MS-DOS
: SOFTWARE: Patent In Release #1.0, Version #1.25
: CURRENT APPLICATION DATA:
: APPLICATION NUMBER: US/08/164,839
: FILING DATE:
: CLASSIFICATION: 435
: PRIOR APPLICATION DATA:
: APPLICATION NUMBER: 08/004,729
: FILING DATE: 14-JAN-1993
: ATTORNEY/AGENT INFORMATION:
: NAME: Oblon, No. 5514573man F.
: REGISTRATION NUMBER: 24,618
: REFERENCE/DOCKET NUMBER: 10-599-0
: TELECOMMUNICATION INFORMATION:
: TELEPHONE: (703)412-3000
: TELEFAX: (703)413-2220
: TELEX: 248855 OPAT UR
: INFORMATION FOR SEQ ID NO: 9:
: SEQUENCE CHARACTERISTICS:
: LENGTH: 2921 base pairs
: TYPE: nucleic acid
: STRANDEDNESS: double
: TOPOLOGY: linear
: MOLECULE TYPE: cDNA to mRNA
: ORIGINAL SOURCE:
: ORGANISM: Theragra chalcogramma
: TISSUE TYPE: liver
: FEATURE:
: NAME/KEY: CDS
: LOCATION: 32..2122
: US-08-164-839-9

alignment_scores:
      Quality: 91.50      Length: 250
      Ratio: 0.855      Gaps: 10
      Percent Similarity: 42.800      Percent Identity: 21.600

alignment_block:
US-09-805-919-1 x US-08-164-839-9/rev ..
Align seg 1/1 to reverse of: US-08-164-839-9 from: 1 to: 2921
      86 SerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysTh 102
      |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      936 TCTGTGTCATGGGCTGACTCAAGTTGGTGATGACGAGCGTGTGGGG... 892
      102 rValPheProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnIleP 119
      ::|||::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      891 ...ATGCCAGGCGATCGCAGCAGCTGTGCAGGCCACGGCTGCAACACACC 846
      119 heGluTyrValArgValAsnAsnAsp..... 127
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
      845 AGCATGGCCATCTTCACTCCAGACTTTGATTGGTCCACTTCTCCAGG 796
```

```

128 .....IleAsnTyrValAsnCysLeuAlaProIleAsnSe 139
      |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
795 ATGGGCACACTGCTCATCCATTGGTTGGCTGACTCCACCTGTGTAAG 746
: APPLICANT: NAKANISHI, KAZUO
: APPLICANT: MOTOKI, MASAO
: APPLICANT: NAGASE, KAZUO
: APPLICANT: MATSUI, HIROSHI
: TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
: NUMBER OF SEQUENCES: 72
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible

156 ertTyrLeuValIle.....Arg 161
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
698 CCATGGCAACGACTGCCCTGCTGATGACAGAGGTCCTCCATCGCTGACGC 649
162 GluIleLeuGlnAlaSerAspLeuAlaGly.....GI 172
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
648 ATGTCGTTGGGGGCTGTGACCTTGCTGCTTAAAGCGGTCAGGATCTC 599
172 ycysArgValGluThrValAlaTrpSerSerAlaProGlyIleSerSera 189
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
598 AAAGCAGATGTCATCACATTTGCTTCAAACTGTCCGAAATTCATGCA 549
189 snLys..... 190
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
548 ACCCACTCATCAATCCGCACTCCCACTGAAATGCGACCATCTCGTTC 499
191 .....SerSer.....ThrLeuSerSe 196
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
498 ATAATGATTCTTGAGCAGACACTTTCATCAGGAGGTACACACATCATC 449
196 rThrHisGlnGlyLeuAlaTyrGlyPheGluLeuSerTrpIysArgAsnL 213
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
448 TTGCACCAAGGATTAACAGCAAGTACATTG.....GTAAGTCCG 405
213 euleuCysArgAsnCys..... 218
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 TCTCTCCAGAGTTATTCCCTGACATCAGTTCAACAGCAAGCGGTAC 355
219 .....AsparSerArgGlyGlyGlyCysThrIleGluGluAsnSerAs 233
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
354 TCGCCAACGCGCAGCAGGCGCAGGAGCTGACACAGTACAGCAAGTATTC 305
233 pArgAlaThrCysArgTyrTrpCysIysGluAspIleHisValSerLysL 250
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
304 CCTCGCGCTCTGC.....TGGTGAACCAACCACTGTTCACCAACACCCTC 261
250 eutThrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePhe 266
      ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 CAGCCGTTGCTGTGATCAACAATATCATCGTTGTTCTTATCCTGCTTT 211

seq_name: /cgn2_6/ptodata/2/lna/5A_COMB.seq:US-08-583-799-9

seq_documentation_block:
: Sequence 9, Application US/08583799
: Patent No. 5607849
: GENERAL INFORMATION:
: APPLICANT: YASUEDA, HISASHI
: APPLICANT: NAKANISHI, KAZUO
: APPLICANT: MOTOKI, MASAO
: APPLICANT: NAGASE, KAZUO
: APPLICANT: MATSUI, HIROSHI
: TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
: NUMBER OF SEQUENCES: 72
: CORRESPONDENCE ADDRESS:
: ADDRESSEE: OBLON, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
: ADDRESSEE: P.C.
: STREET: 1755 Jefferson Davis Highway, Fourth Floor
: CITY: Arlington
: STATE: Virginia
: COUNTRY: U.S.A.
: ZIP: 22202
:
: COMPUTER READABLE FORM:
: MEDIUM TYPE: Floppy disk
: COMPUTER: IBM PC compatible
```



```

; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/583,799
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/004,729
; FILING DATE: 14-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5607849man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-599-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)412-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 9:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA to mRNA
; ORIGINAL SOURCE:
; ORGANISM: Theragra chalcogramma
; TISSUE TYPE: liver
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 32..2122
; US-08-583-799-9

```

```

alignment_scores:
  quality: 91.50      length: 250
  ratio: 0.855      gaps: 10
  percent similarity: 42.800  percent identity: 21.600

```

alignment_block:

US-09-805-919-1 x US-08-583-799-9/rev ..

Align seg 1/1 to reverse of: US-08-583-799-9 from: 1 to: 2921

```

86 SerIleArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThr 102
   |||::: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
936 TCTGTGTCATGGGCTGACTCAAGTTGGTGTGATGCAGCGGTGGGG... 892

102 rValpHeProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnIleP 119
   ::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
891 ...ATGCCAGGCATGCGCAGCACTGTGCAGGCCACGGCTGCAACACCCC 846

119 heGluTrpValArgValAsnAsp..... 127
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
845 AGCATTGGCCATCTTCACTCCAGACTTTGATTGCTCCACTTCTCCAGG 796

128 .....IleAsnTyrValAsnCysLeuAlaProIleAsnse 139
   |||::: |||::: |||::: |||::: |||::: |||::: |||::: |||:::
795 ATGGGCACACTGCTCATCTATTGGTTGGCTGTACTCCACCTGTGTAAAG 746

139 rSerGlnTyrIleProThrSerPheCysSerLysAsnSerThrGlyPheS 156
   ::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
745 TTCTGCATTTCCCAACCAAG...ACTCACCGTCATCGTTGGCAATCA 699

156 eTyrLeuValIle.....Arg 161
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
698 CCAATGGCAACGACTGCCCTGCTGATGTAGACAGGGTCCCATCGCTGAGC 649

162 GluIleLeuGlnAlaSerAspLeuAlaGly.....G1 172
   ::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
648 ATGTGCTTTGGGGGTCTGACCTTGTGCTTAAGCGGTCAAGGATCTC 599

172 yCysArgValGluThrValAlaTrpSerSerAlaProGlyIleSerSerA 189
   ::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||

```

```

seq_name: /cgn2_6/plodata/2/lna/5A_COMB.seq:US-08-344-536-3

seq_documentation_block:
; Sequence 3, Application US/08344536
; Patent No. 5674735
; GENERAL INFORMATION:
; APPLICANT: Onions, David E.
; TITLE OF INVENTION: EHV-4 GLYCOPROTEIN VACCINE
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Bell, Seltzer, Park & Gibson
; STREET: Post Office Drawer 34009
; CITY: Charlotte
; STATE: No. 5674735th Carolina
; COUNTRY: USA
; ZIP: 28234
; COMPUTER READABLE FORM:
; MEDIUM TYPE: floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/344,536
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/961,672
; FILING DATE: 05-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Sibley, Kenneth D.
; REGISTRATION NUMBER: 31,665
; REFERENCE/DOCKET NUMBER: 1749-106
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (919) 881-3140
; TELEFAX: (919) 881-3175
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)

```

```

598 AAAGCAGATGTCATCATCTGCTTCAAACTGTCCGAATTCATGCGCA 549
189 snlys..... 190
548 ACCCACATCATCCGACTCCAGTAGAAATGCGACCATCTCTGCTC 499
191 .....SerSer.....ThrLeuSer 196
498 ATATGATATTCCTTGAGCAGACTTTTCATCAGGAGGTACACAGCATCATC 449
196 rThrHisGlnGlyLeuAlaTyrGlyPheGluLeuSerTrpLysArgAsnL 213
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
448 TTTCACACAGGATTAACAGCAGATACATTTTG.....GTAAGTCCG 405

213 euleuCysArgAsnCys..... 218
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
404 TCCTCTCCAGAAGTTTATTCCTGCATCATCAGTTCAACAGCCAGCGGTAC 355

219 .....AspArgSerArgGlyGlyGlyCysThrIleGluGluAsnSerAs 233
   ::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||:::|||
354 TCGCCAACGGCAGCAGCGGCGAGACTGTACACAGATCAGCAGTAATTCGTT 305

233 pArgAlaThrCysArgTyrTrpCysLysGluAspIleHisValSerLysL 250
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
304 CCTGCGCTCTGC.....TGGTGTAAACACACCACTGTGCACAGACCCCTC 261

250 euthrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePhe 266
   ::||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
260 CAGCCGTCGCTGTCATCACAATATCATCGTGTCTTATTCCTGCTTT 211

```



```
;
; LENGTH: 1560 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 52..1506
; US-09-805-919-1 x US-08-920-562-3

alignment_scores:
    Quality: 90.50      Length: 373
    Ratio: 0.555        Gaps: 16
    Percent Similarity: 43.700    Percent Identity: 18.231

alignment_block:
US-09-805-919-1 x US-08-920-562-3 ..

Align seg 1/1 to: US-08-920-562-3 from: 1 to: 1560

11 ValLeuLeuLeuValThrLeuIleSerSerLysAlaGlnAspIleSe 27
   ::::::::::: ||| ::::::::::: ::::
511 'CTGCTGTATTTCACCGCACATCTTAGCGGAATCTCGAAGGTTGG 560
   ::::::::::: ::::::::::: ::::
27 rGln.....CysValProSerSerCysGlyAspIleGlnIle..... 39
   :::: ::::::::::: ::::
561 CCAGCTGGCGGTATTCCAGACAGCGCTAAGCGTCAACTGTTCAACC 610
   ::::::::::: ::::::::::: ::::
40 .....LysPheProPheArgLeuArgThr 47
   ::::::::::: ::::::::::: ::::
611 TCCCGGCCACACGACGCGTGTACAAATTTCCACTAAACATAAATCT 660
   ::::::::::: ::::::::::: ::::
48 ...AspProGluHisCysGlyArgGlyThrGluLeuAspCysGlnAs 63
   ||| :::: :::: ::::
661 ATAGACTGGCGTACCGGGGAGTTATGTGTGTTACTTGTTCGCAAAA 710
   ::::::::::: ::::::::::: ::::
63 nasGln.....ThrValPheAsnTyrTyrSera 73
   ::::::::::: ::::::::::: ::::
711 CGGCTACTCATTAACAGTACCGGTTACCGTGTAACTACGACGAC 760
   ::::::::::: ::::::::::: ::::
73 rGlePheAspValGlnGluIleAsnTyrArgSerTyrSerIleArgLeu 89
   ::::::::::: ::::::::::: ::::
761 CCTAATGGACCTCTC..... 777
   ::::::::::: ::::::::::: ::::
90 LeuAspProGlyLeuAsnAspGlnArgGlnAsnCysThrValPheProAs 106
   :::: ::::::::::: ::::
778 GTTCAACCAAGTTGAAGGTGA.....AA 803
   ::::::::::: ::::::::::: ::::
106 nHisArg.....AlaSerTyrAspAlaMetThrSerGlnI 118
   ||||| :::: ::::
804 CCACAGAGCCGTGTGCTAGTGTGCTAGCTACTTCCACAACTGTGTA 853
   ::::::::::: ::::::::::: ::::
118 IepheGluTrpValArgValAsnAsnAspIleAsnTyrValAsnCysLeu 134
   ::::::::::: ::::::::::: ::::
854 AGCTGAGGTGTATAAACGCCAAGAGGTGATTTACAAAGTATGTT 903
   ::::::::::: ::::::::::: ::::
135 AlaProIleAsnSer.....SerGlnTyrIleProThrSerPheC 148
   :::: ::::::::::: ::::
904 ACCAATGCTTCTAGTGTGTGGGTGATGTCTCATCACTGCATCTGCAC 953
   ::::::::::: ::::::::::: ::::
140 .....SerGlnTyrIleProThrSerPheC 148
   ::::::::::: ::::::::::: ::::
954 TGTATCAATCCAGCTGACCCCGACGAGAATATCCCCCAGCCCTCGCT 1003
   ::::::::::: ::::::::::: ::::
148 ySer.....LysAsnSerThrGlyPheSer..... 156
   ||||| ::::::::::: ::::
1004 GTAGCATAGATGTACAGAGAGGATATCTTTCTTCGCATGGCCAA 1053
   ::::::::::: ::::::::::: ::::
157 .....TyrLeuValIleArgGluIleLeuGlnAlaSe 167
   ::::::::::: ::::::::::: ::::
1054 GCAGGACGCCCTCTGTGTGTGTGGCCCAACCGTGTCCGTAACGTTGA 1103
   ::::::::::: ::::::::::: ::::
```

```
167 rAspLeuAlaGlyGlyCysArgValGluThr..... 177
   :::: :::: ::::
1104 AGATGTCGACGAGTGTGTGACGAGATGTGTACCTAGCACGAGAGTGT 1153
   ::::::::::: ::::::::::: ::::
178 ..ValAlaTrp.....SerSerAlaProGlyIleSerSerAsnLys 190
   ::::::::::: :::: ::::
1154 TTGTATCGTGGGTGCTTAACGACCATTTACCGGGGTCCCATCACAGAC 1203
   ::::::::::: :::: ::::
191 SerSerThr...LeuSerSerThrHisGlnGlyLeuAlaTyrGlyPheG 206
   :::: :::: ::::
1204 GTACACACGGGAGTTGCTCTCAACCCAGCATTAAGTC..... 1242
   ::::::::::: ::::::::::: ::::
206 uLeuSerTrpLysArgAsnLeuLeuCysArgAsnCysAspArgSerArg 223
   ::::::::::: ::::::::::: ::::
1243 .....AACATGCCGA 1252
   ::::::::::: ::::::::::: ::::
223 IyGlyGlyCysThrIleGluGlnAsnSerAspArgAlaThrCysArgTyr 239
   ::::::::::: ::::::::::: ::::
1253 GTAGCAGGCCCTGTGCGAAGAAACGAGAGCGA..... 1287
   ::::::::::: ::::::::::: ::::
240 TrpCysLysGluAspIleHisValSerLysLeuThrPheArgCysLysVa 256
   :::: :::: ::::
1288 .....GAGTATAACTGCATCAT 1304
   ::::::::::: ::::::::::: ::::
256 IgluTyrTyrSerValTyrValLeuPheGlyGly..... 268
   :::: :::: ::::
1305 'AGAGGTTACCCGACGCGCTTCCAATGTTTCTGACAGCGTTGTATATAG 1354
   :::: :::: ::::
269 .....IleGlyIleGlyGly 273
   ::::::::::: ::::::::::: ::::
1355 ATGATCCCTATTTGTGAGACATGCCGTTTAACTGGCATCATCGCC 1404
   ::::::::::: ::::::::::: ::::
274 ValLeuAla***ArgPheLeuLeuGlyIleProIleLeuIleAlaAlaVa 290
   ||| :::: ::::
1405 GTTACTTGGCGGGCCGACGCGCTAGCGCTGTTGTTCTCATTTACAGCCGT 1454
   ::::::::::: ::::::::::: ::::
290 lValTrpGlnCysLysArg 296
   :::: ::::
1455 TTGTTTACTGCTCAAAA 1473
   ::::::::::: ::::::::::: ::::

seq_name: /cgn2_6/ptodata/2/ina/5A_COMB.seq:US-08-164-839-28

seq_documentation_block:
; Sequence 28, Application US/08164839
; Patent No. 5514573
; GENERAL INFORMATION:
; APPLICANT: YASUEDA, HISASHI
; APPLICANT: NAKANISHI, KAZUO
; APPLICANT: MOTOKI, MASAO
; APPLICANT: NAGASE, KAZUO
; APPLICANT: MATSUI, HIROSHI
; TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED
; FROM FISH
; NUMBER OF SEQUENCES: 72
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: OBION, SPIVAK, MCCLELLAND, MAIER & NEUSTADT,
; P.C.
; STREET: 1755 Jefferson Davis Highway, Fourth Floor
; CITY: Arlington
; STATE: Virginia
; COUNTRY: U.S.A.
; ZIP: 22202
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/164,839
; FILING DATE:
; CLASSIFICATION: 435
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/004,729
```

```

; FILING DATE: 14-JAN-1993
; ATTORNEY/AGENT INFORMATION:
; NAME: Oblon, No. 5514573man F.
; REGISTRATION NUMBER: 24,618
; REFERENCE/DOCKET NUMBER: 10-599-0
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703)412-3000
; TELEFAX: (703)413-2220
; TELEX: 248855 OPAT UR
; INFORMATION FOR SEQ ID NO: 28:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1921 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: unknown
; TOPOLOGY: unknown
; MOLECULE TYPE: DNA (genomic)
; ORIGINAL SOURCE:
; ORGANISM: Theragra chalcogramma
; TISSUE TYPE: muscle
; US-08-164-839-28

```

```

alignment_scores:
      Quality: 89.50      Length: 231
      Ratio: 0.904      Gaps: 8
Percent Similarity: 42.857      Percent Identity: 20.779

```

alignment_block:

US-09-805-919-1 x US-08-164-839-28/rev ..

Align seg 1/1 to reverse of: US-08-164-839-28 from: 1 to: 1921

```

105 ProAsnHisArgAlaSerTyrAspAlaMetThrSerGlnIlePheGluTr 121
    |||::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
857 CCCAGGCATCGCATCTGACGTCAGGCCACGGCTGCAAAACCCAGCATG 808
121 PValArgValAsnAsp..... 127
807 GCCATATTTCACCTCCAGACTTGATTGCTCCACTCTTCAGATGGCA 758
128 .....IleAsnTyrValAsnCysLeuAlaProIleAsnSerSerGln 141
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
757 CACTGCTCATCTTTGGTTGGCTGTACTCCACCTGTGTAAAGTTCCCTGC 708
142 TyrIleProThrSerPheCysSerLysAsnSerThrGlyPheSerTyrLe 158
    ::|||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
707 CATGACCCACACGACTCCACCGTCATCATCGTTGCATTACCATGGC 658
158 uValIleArgGluLe.....L 164
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
657 AACGACTGCCCTGCTGATGATGAGACAGGTCCTCCATCGCTGAAGCATG 608
164 euGlnAlaSerAspLeuAlaGly.....GlyCysArg 174
    ||::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
607 TTGGGGGGTCTGACCTCGCTGGGTTAAAGCGGTCAAGGATCTCAAGACG 558
175 ValGluThrValAlaTyrSerSerAlaPro..... 184
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
557 ATGTCCATTCACATTATCTTCAAACTGTCCGAATTCATGCAACCCACT 508
184 ..... 184
507 CATCAATCCGCACTCCAGTGAATAATGCGACATTCTCGTTCATAATGT 458
185 ..GlyIleSerSerAsnLysSerSer.....ThrLeuSerSerThrHis 198
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
457 ATTCCTTGAGCAGACACTCATCAGGAGGTACACGCGCATCTTCTTTGCAC 408
199 GlnGlyLeuAlaTyrGlyPheGluLeuSerTyrLysArgAsnLeuLeuCy 215
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
407 CAGGATTAAACAGCAAGTACATTTTG.....GTAAAGTCCGTCCTCTC 364
215 sArgAsnCys.....A 219

```

```

363 CAGAAGTTTATTCCTGACATCACTTCAACAGCCACGGGTACTCGCCAA 314
    |||::|
219 spArgSerArgGlyGlyGluCysThrIleGluIleAsnSerAspArgAla 235
    ::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
313 CGCAGACAGGGCGGCGGACTGTACACAGTACAGCAGTAACTGTTGTCGCG 264
236 ThrCysArgTyrTrpCysLysGluAspIleHisValSerLysLeuThrPh 252
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
263 CTCTGC.....TGGTGAACCACTTGTACACGACCCCTCCAGTGT 220
252 eArgCysLysValGluTyrTyrSerValTyrValLeuPhePhe 266
    |||::|::|::|::|::|::|::|::|::|::|::|::|::|::|
219 TCGTGTCTGATCACAATATCATCGTTGATTATTATCCTGCTTT 177

```

seq_name: /cgn2_6/prodata/2/lna/5A_COMB.seq:US-08-583-799-28

seq_documentation_block:

Sequence 28, Application US/08583799

Patent No. 5607849

GENERAL INFORMATION:

APPLICANT: YASUEDA, HISASHI

APPLICANT: MAKANISHI, KAZUO

APPLICANT: MOTOKI, MASAO

APPLICANT: NAGASE, KAZUO

APPLICANT: MATSUI, HIROSHI

TITLE OF INVENTION: GENE ENCODING TRANSGLUTAMINASE DERIVED

NUMBER OF SEQUENCES: 72

CORRESPONDENCE ADDRESS:

ADDRESSEE: OBLON, SPIVAK, MCLELLAND, MAIER & NEUSTADT,

ADDRESSER: P.C.

STREET: 1755 Jefferson Davis Highway, Fourth Floor

CITY: Arlington

STATE: Virginia

COUNTRY: U.S.A.

ZIP: 22202

COMPUTER READABLE FORM:

MEDIUM TYPE: Floppy disk

COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS

SOFTWARE: PatentIn Release #1.0, Version #1.25

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/583,799

FILING DATE:

CLASSIFICATION: 435

Prior APPLICATION DATA:

APPLICATION NUMBER: 08/004,729

FILING DATE: 14-JAN-1993

ATTORNEY/AGENT INFORMATION:

NAME: Oblon, No. 5607849man F.

REGISTRATION NUMBER: 24,618

REFERENCE/DOCKET NUMBER: 10-599-0

TELECOMMUNICATION INFORMATION:

TELEPHONE: (703)412-3000

TELEFAX: (703)413-2220

TELEX: 248855 OPAT UR

INFORMATION FOR SEQ ID NO: 28:

SEQUENCE CHARACTERISTICS:

LENGTH: 1921 base pairs

TYPE: nucleic acid

STRANDEDNESS: unknown

TOPOLOGY: unknown

MOLECULE TYPE: DNA (genomic)

ORIGINAL SOURCE:

ORGANISM: Theragra chalcogramma

TISSUE TYPE: muscle

```

alignment_scores:
      Quality: 89.50      Length: 231
      Ratio: 0.904      Gaps: 8

```


OM of: US-09-805-919-1 to: EST:* out_format : pfs
Date: Jun 4, 2002 10:45 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODEL=frame+pn.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09805919/runat_04062002.121211_7623/app_query.fasta_1.369
-DB=EST -QFMT=fastap -SUFFIX=rst -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPCL=0.000 -LOOPEXT=0.000 -GAPOP=4.500
-GAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPEXT=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blsum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09805919 @CGN1_1_6102
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPXY -WAIT -THREADS=1

Search information block:

Query: US-09-805-919-1
Query length: 308
Database: EST.*
Database sequences: 13736207
Database length: -1841457050
Search time (sec): 1592.490000

Sequence	Strd	Orig	2Score	EScore	Len	! Documentation
gb_est1:AI775018	+	705.00	1352.90	4.1e-66	538	! AI775018 EST256118 tomato resist
gb_est2:BI434492	+	697.50	1336.27	3.5e-65	639	! BI434492 EST537253 P. infestans
gb_est1:AI897599	+	624.50	1196.59	2.1e-57	533	! AI897599 EST267042 tomato ovary
gb_est1:AI490332	+	516.50	989.23	7.5e-46	432	! AI490332 EST248658 tomato ovary
gb_est2:BI179642	+	294.50	551.80	1.7e-21	717	! BI179642 EST520587 CSTE Solanum
gb_est2:BE342777	+	288.00	541.08	6.8e-21	611	! BE342777 EST395621 potato stold
gb_est2:BI922602	+	274.50	512.81	2.6e-19	725	! BI922602 EST542506 tomato callu
gb_est2:BI922851	+	266.00	497.39	1.9e-18	661	! BI922851 EST542755 tomato callu
gb_est2:BG136204	-	262.50	491.01	4.2e-18	638	! BG136204 EST476646 wild tomato
gb_est1:AM737395	+	258.00	482.52	1.2e-17	625	! AM737395 EST338738 tomato flowe
gb_est2:BI179655	+	250.00	468.01	8.0e-17	573	! BI179655 EST520600 CSTE Solanum
gb_est2:BG455609	+	234.50	436.29	4.7e-15	655	! BG455609 NF061D05PL1F1044 Phosf
gb_est1:AM560216	+	233.50	436.50	4.6e-15	547	! AM560216 EST315264 DSIR Medicag
gb_est1:AM906806	+	233.00	435.62	5.1e-15	543	! AM906806 EST342929 potato stold
gb_est2:BI922909	+	232.00	430.47	9.9e-15	710	! BI922909 EST542813 tomato callu
gb_est2:BF637888	+	227.50	423.25	2.5e-14	625	! BF637888 NF030B07PL1F1059 Phosf
gb_est2:BE920840	+	226.50	422.08	2.9e-14	586	! BE920840 EST424609 potato leave
gb_est2:BF645984	+	226.00	419.79	3.9e-14	654	! BF645984 NF065E11EC1F1086 Elic
gb_est2:BE923898	+	224.00	419.98	3.8e-14	465	! BE923898 EST427667 potato leave
gb_est2:BF643246	+	223.50	417.09	5.5e-14	546	! BF643246 NF003C11EC1F1083 Elic
gb_est1:AM906805	+	223.00	418.89	4.4e-14	433	! AM906805 EST342928 potato stold
gb_est2:BI434233	+	220.50	408.38	1.7e-13	695	! BI434233 EST536994 P. infestans
gb_est1:AI484361	+	219.00	407.26	1.9e-13	598	! AI484361 EST249009 tomato suscep
gb_est2:BE472530	+	214.50	400.56	4.6e-13	504	! BE472530 EST417383 potato stold
gb_est1:AM694208	+	213.50	395.43	8.9e-13	658	! AM694208 NF073D09ST1F1077 Devel
gb_est2:BF633219	+	210.50	392.93	1.2e-12	498	! BF633219 NF054H02DT1F1026 Droug
gb_est1:AM981159	-	210.00	387.24	2.5e-12	739	! AM981159 EST392353 DSIL Medicag
gb_est2:BE919964	+	203.50	379.33	7.0e-12	498	! BE919964 EST423733 potato leave
gb_est1:BE191568	+	195.50	363.78	5.1e-11	498	! BE191568 sn78b04.yl Gm-c1038 gl
gb_est1:AM695109	+	195.00	360.52	7.8e-11	603	! AM695109 NF091D01ST1F1012 Devel
gb_est1:AI772810	+	187.00	343.42	7.0e-10	687	! AI772810 EST244554 potato stold
gb_est2:BF188267	+	184.00	343.35	7.0e-10	424	! BF188267 EST44554 potato stold
gb_est1:AM907203	+	184.00	343.05	7.3e-10	435	! AM907203 EST343326 potato stold
gb_gss:AO273669	-	178.50	328.40	4.8e-09	606	! AO273669 ndxb0030P03r CUGI Rice
gb_est1:AM218767	+	175.50	321.46	1.2e-08	665	! AM218767 EST301247 tomato root
gb_est2:BG096676	+	174.00	324.95	7.5e-09	389	! BG096676 EST461195 potato leave
gb_gss:BH022645	+	168.50	308.01	6.6e-08	657	! BH022645 GH_MBB0003G06r Gossyp
gb_est2:BG448627	+	165.50	302.01	1.4e-07	666	! BG448627 NF044E04NRI1F1024 Nodul
gb_est1:AM208085	+	157.50	292.59	4.7e-07	399	! AM208085 M111118e DSIR Medicag
gb_est1:AV938087	+	154.50	280.82	2.1e-06	656	! AV938087 AV938087 K. Sato unpub

gb_est1:AM688209 + 152.00 276.16 3.9e-06 645 | AM688209 NF004G02ST1F1000 |
gb_est1:AM685490 + 152.00 275.89 4.0e-06 660 | AM685490 NF030F02NRI1F1000 |
gb_est2:BG599159 + 148.50 271.71 6.9e-06 530 | BG599159 EST504059 CSTS So.
gb_est1:AM617540 + 148.00 270.19 8.4e-06 555 | AM617540 EST323951 L. hirs
gb_est1:AM617826 + 148.00 269.62 9.0e-06 582 | AM617826 EST324225 L. hirs

seq_name: gb_est1:AI775018

seq_documentation_block:

LOCUS AI775018 538 bp mRNA linear EST 18-MAY-2001
DEFINITION EST256118 tomato resistant, Cornell Lycopersicon esculentum CDNA
clone CLER14G5, mRNA sequence.

ACCESSION AI775018 GI:5273059

VERSION AI775018.1 GI:5273059

KEYWORDS EST.

SOURCE tomato.

ORGANISM Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asterales; euasterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.
1 (bases 1 to 538)

REFERENCE D' Ascenzo, M., He, X., Lyman, J., Matern, A. L., Vision, T., Holt, I. E.,
Liang, F., Upton, J., Ronning, C. M., Craven, M. B., Fujii, C. Y., Bowman
, C. L., Nierman, W., Fraser, C. M., Venter, J. C., Tanksley, S. D.,
Giovannoni, J. J. and Martin, G. B.
Generation of ESTs from Pseudomonas resistant tomato
Unpublished (1999)

COMMENT Contact: CUGI
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
5 prime sequence.

TITLE JOURNAL

FEATURES source

1. 538
/organism="Lycopersicon esculentum"
/cultivar="R11-12 (35S::Pto in Rio Grande x Money Maker)"
/db_xref="taxon:4081"
/clone="CLER14G5"
/clone_1lb="tomato resistant, Cornell"
/tissue_type="leaf"
/dev_stage="4-week old"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; CLER - Tomato Pseudomonas Resistant EST library.
Directionally cloned cDNAs inserted into pBluescript SK(-)
) at 5' end with EcoRI and 3' end with XhoI site."

BASE COUNT 145 a 126 c 107 g 160 t

ORIGIN

alignment_scores:

Quality: 705.00 Length: 179

Ratio: 4.406 Gaps: 2

Percent Similarity: 89.385 Percent Identity: 74.302

alignment_block:

US-09-805-919-1 x AI775018

Align seg 1/1 to: AI775018 from: 1 to: 538

27 SerGlnCysValProSerSerCysGlyAspLe...GlnIleLysPhepr 42

|||||

2 TCTCAATGTCCTCCATCCCTGTCGATAGTGGCACTAAATTTCC 51

42 opheargleuargThrAspProGluHisCysGlyArgArgGlyTyrGluL 59

|||||

52 TTTCGATGAGGTCTGATCCGACGATGTGTGAAGCTGATATGAAC 101

59 euaspcysGlnAsnAsnGlnThrValPheAsnTyrLysSerArgIlePhe 75

|||||

102 TCGATTGCCAGATAACGAACCAATTTCCATTACAGTCGCGTAATTC 151

76 AspvAlGInGluIleAsnTyrArgSerTyrSerIleArgLeuAspPr 92
|||||
152 TAGTCCAGAGATTACTATACCACTTCTCAATCAGCTACTGATCC 201
92 OGlyLeuAsnAspGlnArgGluAsnCysThrValPheProAsnHisArg 109
|||||
202 CAGCCTGAAAAATCAGACCCGAAATGCTCTGTGTTTCCACACACAGCG 251
109 lAserTyrAspAlaMetThrSerGlnIlePheGluTrpValArgValAsn 125
|||||
252 CAAGTTACGATGCCATGACTAGCCATACTTCGATGCGCTGCTGAT 301
126 AsnAspIleAsnTyrValAsnCysLeuAlaProIleAsnSerSerGlnTyr 142
|||||
302 AACGATATCAATTATGTCACTGTCCGGGTCAATCACTCCTCAGTATA 351
142 rIleProThrSerPheCysSer...LysAsnSerThrGlyPheSerTyrL 158
|||||
352 CATTCCTACAGATTTTGTAGCACAGCTACAAATTCAGTTTGTAGCTATC 401
158 euValIleArgGluIleLeuGlnAlaSerAspLeuAlaGlyGlyCysArg 174
|||||
402 TTGTCGTAGGTGAATATTGCAAGCTTCAGATTAGCAGTCGGTTGCAGA 451
175 ValGluThrValAlaTrpSerSerAlaProGlyIleSerSerAsnLysSe 191
|||||
452 GTGAAACTGTGGCATGAGAGTGCAGCTCCGGGATTTCCACTTAACGTATC 501
191 rSerThrLeuSerSerThrHisGlnGlyLeuAlaTyr 203
|||||
502 GTCTTCGTTATCCAGCATTCACCAAGCTCTGGGTTAC 538

seq_name: gb_est2:BI434492

seq_documentation_block:

LOCUS BI434492 639 bp mRNA linear EST 21-AUG-2001
DEFINITION EST537253 P. infestans-challenged leaf Solanum tuberosum cDNA clone
PPCBR09 5' sequence, mRNA sequence.

ACCESSION BI434492

VERSION BI434492.1 GI:15259182

KEYWORDS EST.

SOURCE potato.

ORGANISM Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE 1 (bases 1 to 639)

AUTHORS Restrepo,S., Griffiths,H.M., Smart,C.D., Cho,J., Chlemingo,A.,

Bougril,O., Buell,C.R., Ronning,C.M., Fry,W.E. and Baker,B.

Generation of ESTs from Potato Leaves Challenged with Phytophthora

infestans, Compatible Interaction

Unpublished (2000)

JOURNAL Contact: Cathy Ronning

COMMENT The Institute for Genomic Research

For clone info: please contact Research Genetics, Libraries

Division tel 1-800-711-6195, email cdaetresgen.com

Seq primer: M13F-R.

FEATURES Location/Qualifiers

source

1..639

/organism="Solanum tuberosum"

/cultivar="Kennebec"

/db_xref="taxon:4113"

/clone="PPCBR09"

/clone_1lb="P. infestans-challenged leaf"

/tissue_type="leaf"

/dev_stage="6 week old"

/lab_host="SOLR"

/note="Vector: pBluescript SK(-): Site_1: EcoRI; Site_2:

XhoI; Whole plants were challenged with 450,000

sporangia/ml P. infestans US-1(US 940501) in Biotron

(Madison, Wisconsin). Leaf tissue was collected at 1, 2,

5, 12, and 24 hours post-challenge and frozen in liquid

nitrogen immediately upon removal. Kennebec plants showed
no signs of HR. Kathadin plants (susceptible to P.
infestans US-1) were used as controls and showed
infection. NOTE: We cannot exclude the possibility that
this sequence is actually derived from Phytophthora rather
than potato."

BASE COUNT 200 a 144 c 101 g 194 t
ORIGIN

alignment_scores:

Quality: 697.50 Length: 197

Ratio: 4.079 Gaps: 4

Percent Similarity: 86.802 Percent Identity: 69.543

alignment_block:

US-09-805-919-1 x BI434492 ..

Align seg 1/1 to: BI434492 from: 1 to: 639

8 PheAlaCysValLeuLeuLeuValThrLeuIleSerSerTyrAlaGly 24
|||||
47 TTTCCTAGTGTGTTTCTCCTGTTTCATTCATTAACAAGCTCTATAGCCCA 96
24 nAsp...IleSerGlnCysValProSerSerCysGlyAspIleGln...I 39
|||||
97 AAATAGTACGTCTCAATGTGCCCTTCTCTCGTGCGGATATTCGCATA 146
39 lElySPheProPheArgLeuArgThrAspProGluHisCysGlyArgArg 55
|||||
147 TAAATTCCTCCCTCCGATGAGGACTGATCCGAGCATTTGTGGCAAAACAT 196
56 GlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPheAsnTyrLysSe 72
|||||
197 GAATACGAAGCTGATGGCAGACAAATCAAACTCTTACACTTACAGTGC 246
72 rArgIlePheAspValGlnIleAsnTyrArgSerTyrSerIleArgL 89
|||||
247 CAGAAATCTCTACGTGCAGAAATTAACACAAAGCTACATGATAGGC 296
89 euLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPhePro 105
|||||
297 TACTCGATCCTGGCCTAAAGATCAGAACGAAATTCCTCTGTTTCCT 346
106 AsnHisArgAlaSer...TyrAspAlaMetThrSerGlnIlePheGluTr 121
|||||
347 GATTACAGACGAATATTATACGAGGCTTAACCTAGTGAATTTCCAAAT 396
121 pValArgValAsnAsnAspIleAsnTyrValAsnCysLeuAlaProIleA 138
|||||
397 GATTCAATATTACAAATAGCACTAATGTCACTGTGGAACCTCCATCA 446
138 snSerSerGlnTyrIleProThrSerPheCysSerLysAsnSerThr... 153
|||||
447 ACTCGTCACAAATACATTCACAGCACTTTTGTAGAACAAATACAACTTCG 496
154GlyPheSerTyrLeuValIleArgGluIleLeuGlnAlaSe 167
|||||
497 CCCAAGGCCAAATTTTAGTTATCTTGTCTTAAAGAAATATGGCAAGCTTC 546
167 rAspLeuAlaGlyGlyCysArgValGluThrValAlaTrpSerSerAlaP 184
|||||
547 GGATTTGGAAAAACGGTTGCAGAGTTGAAGACTGTGGCGCTGCTTCAGCTC 596
184 roGlyIleSerSerAsnLysSerSerThrLeuSerSerThr 197
|||||
597 CAGGTATTTTACAAACAGATCGCCTTGTGTAGGAGCACT 637

seq_name: gb_est1:AI897599

seq_documentation_block:

LOCUS AI897599

DEFINITION EST267042 tomato ovary, TAMU Lycopersicon esculentum cDNA clone

533 bp mRNA linear EST 18-MAY-2001

```

ACCESSION      CLED28G17, mRNA sequence.
VERSION        AI897599
KEYWORDS       AI897599.1 GI:5603501
SOURCE         tomato.
ORGANISM       Lycopersicon esculentum
AUTHORS        Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
                Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
                Asteridae; euasterids I; Solanales; Solanaceae; Solanum;
                Lycopersicon.
REFERENCE      1 (bases 1 to 533)
TITLE          Alcala,J., Vrebalov,J., White,R., Matern,A.L., Vision,T., Holt,I.E.,
                Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman,
                C.L., Niernan,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley,
                S.D. and Giovannoni,J.
JOURNAL        Generation of ESTs from tomato carpel tissue
COMMENT        Unpublished (1999)
                Contact: CUGI
                Clemson University Genomics Institute
                Clemson University
                100 Jordan Hall, Clemson, SC 29634, USA
                Email: http://www.genome.clemson.edu/orders/index.html
                5 prime sequence.
FEATURES
    source
        1..533
        /organism="Lycopersicon esculentum"
        /cultivar="TA496"
        /db_xref="taxon:4081"
        /clone="CLED28G17"
        /clone_lib="tomato ovary, TAMU"
        /tissue_type="carpel"
        /dev_stage="5 days pre-anthesis to 5 days post-anthesis"
        /lab_host="X11-Blue MRF"
        /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
                XhoI; CLED - Tomato Carpel EST Library. OligodT primed and
                directionally cloned cDNA in vector lambda ZAP II with 5'
                and 3' ends located at the EcoRI and XhoI sites,
                respectively."
BASE COUNT     135 a 103 c 120 g 175 t
ORIGIN
alignment_scores:
    Quality: 624.50      Length: 179
    Ratio: 4.055         Gaps: 2
    Percent Similarity: 86.034    Percent Identity: 66.480
alignment_block:
    US-09-805-919-1 x AI897599 ..
Align seg 1/1 to: AI897599 from: 1 to: 533
113 GlnArGluAsnCysThrValPheProAsnHisArgAlaSerTyrAspAl 113
114 ||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
2 CAGACCGAAATTCCTCTGTTTCCACACACACAGGCGAGTTACGATGC 51
113 aMetThrSerGlnIlePheClnuTrpValArgValAsnAsnAspIleAsnT 130
114 ||||| ||||| ||||| |||||:::|||||:::|||||:::|||||
52 CATGACTAGCCTAATCTTCGCGATGGCTTCGTTGATTGATACGATATCAATT 101
130 yValAsnCysLeuAlaProIleAsnSerSerGlnTyrIleProThrSer 146
114 ||||| ||||| ::||| ||||| ||||| ||||| ||||| |||||
102 ATGTCAACTGTGCGGTCAAAATCAACTCCTCAGTATACATTCCTACAAGT 151
147 PheCysSer...LysAsnSerThrGlyPheSerTyrIleValIleArgG1 162
114 ||||| ||||| ::|||:::|||||:::|||||:::|||||:::||
152 TTTTGTAGCACAGCTACAAATCTAGTTTGTAGCTATCTTGTCTAGGTGA 201
162 uIleLeuGlnAlaSerAspIleuAlaG1yG1yCysArgValGluThrValA 179
114 ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
202 AATATTGCAAGCTTCAGATTACAGATCGGTTGCAGAGTTGAACGTGTGG 251
179 latrpSerAlaProGlyIleSerSerAsnLysSerSerThrIleuSer 195

```

252	CATGAGATGCATCTCCGGGATTTCAACCTAACGATATCGTCTTCGTATCC	301
196	SerThrHisGlnGlyLeuAlaTyrGlyPheGluLeuSerTrpLysArgAs	212
302	AGCATTCACCAAGCTCTGGGTACGATTTGATCTCTCTTGAGGCGTGA	351
212	nleuLeuCysArgAsnCysAspArgSerArgGlyGlyIuGlyThrIleG	229
352	TTTCTTATGATAGAGTGT.....GGAGAGATGATGAATGTCTTTTCA	395
229	IuGluAsnSerAspArgAlaThrCysArgTyrTrpCysLysGluAspIle	245
396	AAGACAATAGTAGTATGACCAACTGTATTCATTACTGTAAAGAGACACT	445
246	HisValSerLysLeuThrPheArgCysLysValGluTyrTrpSerValTy	262
446	CCTGTTTCTGAACGTTCTTTCCGATGCAAAGTCGAGTACTATTCTGTTT	495
262	rValLeuPhePheGlyIleGlyIleGlyIleGlyIleGlyVal	274
496	TGTATCGATCTATGGTGGCATGCAATTGGTGCACTT	532
seq_name: gb_est1:AI490332		
seq_documentation_block:		
LOCUS	AI490332	432 bp mRNA linear EST 18-MAY-2001
DEFINITION	EST248658 tomato ovary, TAMU Lycopersicon esculentum cDNA clone	
ACCESSION	CLD24P14, mRNA sequence.	
VERSION	AI490332	
KEYWORDS	AI490332.1 GI:4385642	
SOURCE	EST.	
ORGANISM	tomato.	
	Lycopersicon esculentum	
	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;	
	Asteridae; euasterids I; Solanales; Solanaceae; Solanum;	
	Lycopersicon.	
REFERENCE	1 (bases 1 to 432)	
AUTHORS	Alcala,J., Vrebalov,J., White,R., Matern,A.L., Viston,T., Holt,I.E.,	
	Liang,F., Upton,J., Ronning,C.M., Craven,M.B., Fujii,C.Y., Bowman	
	,C.L., Nierman,W., Fraser,C.M., Venter,J.C., Martin,G.B., Tanksley	
	,S.D. and Giovannoni,J.	
TITLE	Generation of ESTs from tomato carpel tissue	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: CUGI	
	Clemson University Genomics Institute	
	Clemson University	
	100 Jordan Hall, Clemson, SC 29634, USA	
	Email: http://www.genome.clemson.edu/orders/index.html .	
FEATURES	location/Qualifiers	
Source	1..432	
	/organism="Lycopersicon esculentum"	
	/cultivar="TA496"	
	/db_xref="taxon:4081"	
	/clone="CLD24P14"	
	/clone_lib="tomato ovary, TAMU"	
	/tissue_type="carpel"	
	/dev_stage="5 days pre-anthesis to 5 days post-anthesis"	
	/lab_host="XLI-Blue MRF"	
	/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:	
	XhoI; cLSD - Tomato Carpel EST Library. OligoDT-primed and	
	directionally cloned cDNA in vector Lambda ZAP II with 5'	
	and 3' ends located at the EcoRI and XhoI sites,	
	respectively."	
BASE COUNT	112 a 85 c 99 g 136 t	
ORIGIN		
alignment_scores:		
Quality:	516.50	Length: 145
Ratio:	4.132	Gaps: 2
Percent Similarity:	86.207	Percent Identity: 68.276

alignment_block:
US-09-805-919-1 x AI490332 ..

Align seg 1/1 to: AI490332 from: 1 to: 432

```

 97 GlnArgGluAsnCysThrValPheProAsnHisArgAlaSerTyrAspAl 113
    ||| |||||:::|||||:::|||||:::|||||:::|||||:::|||||
  2 CAGACCGAAATGCTCTGTTTTCACACACAGGCAAGTTACGATGC 51
113 amethrSerGlnIlePheGluTrpValArgValAsnAspIleAsnT 130
    ||||| ||||| ||||| |||||:::|||||:::|||||:::|||||
 52 CATGACTAGCCTAATCTTCGGATGGCTTCGTGTGATTAACGATATCAATT 101
130 yrValAsnCysLeuAlaProIleAsnSerSerGlnTyrIleProthrSer 146
    ||||| |||||::: ||||| ||||| ||||| ||||| ||||| |||||
102 ATGTCAACTGTCCGGGTCAAAATCAACTCCTCAGTATACATTCCACAAGT 151
147 PheCysSer...LysAsnSerThrGlyPheSerTyrLeuValIleArgG1 162
    ||||| |||||::: ||||| ||||| ||||| ||||| ||||| |||||
152 TTTGTAGACAGCTACAAATCTAGTTTGTAGCTATCTTGTGCTAGCTGA 201
162 uIleLeuGlnAlaSerAspLeuAlaGlyGlyCysArgValGluThrValA 179
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
202 AATATTGCAAGCTTCAGATTTAGCAGTCGCGTTGCAGAGTTGAACCTGTGG 251
179 IaTrpSerSerAlaProGlyIleSerSerAsnLysSerSerThrLeuSer 195
    ||||| |||||::: ||||| ||||| ||||| ||||| ||||| |||||
252 CATGAGTGCAGCTCCGGGAGTTTCACCTAACGTAATCGTCTCGTATATCC 301
196 SerThrHisGlnGlyLeuAlaTyrGlyPheGluLeuSerTrpLysArgAs 212
    ||| |||||::: ||||| ||||| ||||| ||||| ||||| |||||
302 AGCATTCACCAAGCTCTGGGTTACGAGATTGATCTCTCTTGAGAGCGCTGA 351
212 nLeuLeuCysArgAsnCysAspArgSerArgGlyGlyGlyCysThrIleG 229
    ::||| |||||::: ||||| ||||| ||||| ||||| ||||| |||||
352 TTCTTATGCAGAGAGTGT.....GGAGAGATGATGAATGTGTTTCA 395
229 luGluAsnSerAspArgAlaThrCysArgTyrTrp 240
    ::||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
396 AAGACAAATAGGGATGTAGCACTTGTATTCATTAC 430
```

seq_name: gb_est2:BI179642

seq_documentation_block:

LOCUS BI179642 717 bp mRNA linear EST 09-JUL-2001
DEFINITION EST520587 CSTE Solanum tuberosum cDNA clone CSTE19J19 5' sequence,
mRNA sequence.

ACCESSION BI179642
VERSION BI179642.1 GI:14645453
KEYWORDS EST.

SOURCE
ORGANISM

potato.
Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.

REFERENCE
AUTHORS 1 (bases 1 to 717)
Van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
Baker,B.

TITLE
JOURNAL Unpublished (2001)
COMMENT Contact: Cathy Ronning
The Institute for Genomic Research
For clone info: please contact Research Genetics, Libraries
Division tel 1-800-711-6195, email cdaeresgen.com

FEATURES
source Location/Qualifiers
1..717

/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone="CSTE19J19"

/clone_lib="CSTE"
/tissue_type="axillary buds of stem explants; growing
sink-tubers"
/dev_stage="7, 8 and 10 days"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue supplied by Christian Bachem and Richard
Visser (Department of Plant Breeding, Wageningen
University, The Netherlands). The csta libraries will
attempt to capture the induction and initiation/initial
growth of the tuber in an in vitro system as described in
Bachem et al. (Plant Journal, 1996). Small microtubers
develop from axillary buds attached to stem explants when
placed on a high sucrose medium (10%). Visible
morphological changes occur synchronously at day five in
the axillary buds. The first library, csta (1-20) consists
of axillary buds harvested on days 1-3. This targets
those genes involved in induction of the microtubers. The
following libraries, csta (21-40) and csta (41-60),
capture genes involved in tuber initiation and outgrowth.
This library is noted as p3 in Tanksley lab notebooks."

alignment_scores:
Quality: 294.50 Length: 210
Ratio: 2.150 Gaps: 7
Percent Similarity: 65.238 Percent Identity: 35.238

alignment_block:
US-09-805-919-1 x BI179642 ..

Align seg 1/1 to: BI179642 from: 1 to: 717

```

 6 leuLeuPheAlaCysValLeuLeuValThrLeuIleSerSerSerly 22
    ||| ::| ::|||::: |||||::: |||||::: |||||::: |||||
 29 CTTAAATACCTTACATTAATCCTTCAATCATCTTCTGCAAAACATCCAA 78
22 salacGlnAspIleSerGlnCysValProSerSerCysGlyAspIleGln. 38
    ::||| |||||::: ||||| ||||| ||||| ||||| ||||| |||||
 79 TGCTTGAAGATC...AAGTGTCTGATTTCTGCTGTGGCATATCCGTA 125
39 ..IleLysPheProPheArgLeuArgThrAspProGluHisCysGlyArg 54
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||
126 ACATAAGACACCCCTTTTCATTTAAACACCGATCCAAACATTTGCATTTT 175
55 ArgGlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPheAsnTyrLy 71
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
176 TCAGGACTTGAAATTAGCTTGTGAAGGTTAACCAACAGTTATATGTTATC 225
71 sSerArgIlePheAspValGlnGluIleAsnTyrArgSerTyrSerIleA 88
    ::|||::: |||||::: |||||::: |||||::: |||||::: |||||
226 CTCCAAGAGAGTTGCACGTTCAAGCATCGACTATGATATATCAATAAATTC 275
88 rGleLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPhe 104
    ::|||::: |||||::: |||||::: |||||::: |||||::: |||||
276 ACCTGATGATCCGACTTTACAAACACCAAGATGATCTATGTTCTCTCATA 325
105 ProAsn.....HisArgAlaSerTyrAspAlaMetThrSerG1 117
    |||||::: |||||::: |||||::: |||||::: |||||::: |||||
326 CCTTCTAGTTTACCTTTCACAAATACAGTAGTTTCTTCAGCGGCTCGTA 375
117 nIlePheGluTrpValArg...ValAsnAsnAspIleAsnTyrValAsnC 133
    ::| ::| ||||| ||||| ||||| ||||| ||||| |||||
376 CTATTATACATCAGATCGTAGTGTAGCAGCGCGGATTTTCATGTTAACT 425
133 ysLeuAlaProIleAsnSerSerGlnTyrIleProThrSerPheCysSer 149
    ||| ::|||::: |||||::: |||||::: |||||::: |||||
426 GTCCATTGCTGTTAATGATTCAACATTTGTGAAATTTAGTGCTGCCAAA 475
150 LysAsnSerThrGlyPheSerTyrLeuValIleArgGluIleGlnAl 166
    ::| ::||| ||||| ||||| ||||| ||||| |||||
```


100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics Institute

FEATURES
Seq primer: T3.
Location/Qualifiers
1..725

/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEC78A19"
/clone_lib="tomato callus"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2: XhoI; supplier: Boyce Thompson Institute; sequencing: The Institute for Genomic Research; CLEC - Cotyledons of seedlings 7-10 days post-germination were excised, cut at both ends and placed on MS medium with no selection. Mixed callus was harvested at 25 and 40 days and included undifferentiated masses. Tomato Callus EST library"

BASE COUNT 205 a 148 c 128 g 244 t
ORIGIN

alignment_scores:
Quality: 274.50 Length: 293
Ratio: 1.654 Gaps: 12
Percent Similarity: 56.655 Percent Identity: 26.962

alignment_block:
US-09-805-919-1 x BI922602 ..

Align seg 1/1 to: BI922602 from: 1 to: 725

```
6 LeuLeuPheAlaCysValLeuLeuValThrLeuIleSerSerSerly 22
   |||  ::  ::::: |||  ::  ::::: |||  ::  ::::: |||  ::
25 CTTAGATATTTACATTAATCCCTCAATCATATTTCTGCAACATCCAA 74
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
22 sAlaGlnAspIleSerGlnCysValProSerSerCysGlyAspIleGln. 38
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
75 TGCTTGAAGATCAAC..TGCTCTGATCTGCTGCTGCGCATATCCGTA 121
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
39 ..IleLysPheProPheArgLeuArgThrAspProGluHisCysGly... 53
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
122 GCATAAAATACCCCTTTTCATTTAAACACTGATCCAAACATTTGCTGTT 171
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
54 ArgArgGlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPheAsnTy 70
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
172 TTTTCAGGATTTGAATTAGCTTGTAAGGTAAACCAACCGTTATATGTT 221
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
70 rLysSerArgIlePheAspValGlnGluIleAsnTyrArgSerTyrSerI 87
   ::::: |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
222 ATCCTCCAAAGAGTTGCATGTTCAAGGCATCGACTATGCTAATAACACA 271
   ::::: |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
87 leArgLeuLeuAspProGlyLeuAsnAspGlnArgGluAsnCysThrVal 103
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
272 TTCATTTGGTAGATCCGACTTTA..CAACACATGATCATGCTCTCTCTC 318
   |||  |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
104 pheProAsn.....HisArgAlaSerTyrAspAlaMetThrSerGlnI 118
   ::::: |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
319 CTACCTTTCTAAGCTTAACCTTCAACAAATACGGTAATTTCTTCTGTGCA 368
   ::::: |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
118 ephe.....GluTrpValArgValAsnAsnAspIleAsnTyrValAsn 133
   ::::: |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
369 CTATTATGATCAGCGGTATAGAGTAGCAGCCCATTTTCATGTCAATT 418
   ::::: |||  |||  |||  |||  |||  |||  |||  |||  |||  |||
133 ysLeuAlaProIleAsnSerSerGlnTyrIleProThrSerPheCysSer 149
   |||  ::  ::::: |||  ::  ::::: |||  ::  ::::: |||  ::
419 GTCCATCTGTGTGATGATTCGACATTTGTGAAATTAGCGGCTGC... 465
   |||  ::  ::::: |||  ::  ::::: |||  ::  ::::: |||  ::
150 LysAsnSerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAl 166
```

```
466 ...AAATCAAGCAATACACTTACCTAAGATTGGAGAA...ATGAAGT 509
   ::|||::: ::::: |||  |||  |||  ::::: |||
166 aserAspLeuAlaGlyGlyCysArgValGluThrValAlaTrpSerSera 183
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
510 CTCTCAAGTAGATGATGATGCAAGCGAGTTTATAGGCTTAACCTCAT 559
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
183 laProGlyIleSerSerAsnLysSerSerThrLeuSerSerThrHisGln 199
   |||  ::|||::: ::::: |||  ::  ::|||  |||  |||  |||
560 GGCCATAATATTAAC.....AACATTTCCTATTTCTAATATTCATCAA 600
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
200 GlyLeuAlaTyrGlyPheGluLeuSerTrpLysArgAsnLeuLeuCysAr 216
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
601 GCAATTCTCTACGGATTGAGTTTCCTTAT..... 630
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
216 gasnCysAspArgSerArgGlyGlyGlyCysThrIleGluIleAsnSera 233
   |||  ::|||::: |||  |||  |||  |||  ::::: |||
631 .....TCTTCGAGTGGG..... 642
   |||  ::|||::: |||  |||  |||  |||  ::::: |||
233 sPArgAlaThrCysArgTyrTrpCysLysGluAspIleHisValSerLys 249
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
642 ..... 642
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
250 LeuThrPheArgCysLysValGluTyrTyrSerValTyrValLeuPhePh 266
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
643 .....TTCTTCGGTTGTTTCCCGGTGATTTG 668
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
266 eglyGlyIleGlyIleGlyGlyValLeuAla**ArgPheLeuLeuGlyI 283
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
669 GGGA.....GCGAAATTGTGATAGGCC 691
   ::|||::: |||  |||  |||  |||  |||  ::::: |||
```

seq_name: gb_est2:BI922851

seq_documentation_block:

LOCUS BI922851 661 bp mRNA linear EST 18-OCT-2001

DEFINITION EST542755 tomato callus Lycopersicon esculentum cDNA clone

ACCESSION BI922851

VERSION BI922851.1 GI:16221050

KEYWORDS EST.

SOURCE

ORGANISM

Lycopersicon esculentum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

Asteridae; easterids I; Solanales; Solanaceae; Solanum;

Lycopersicon.

1 (bases 1 to 661)

Alcala,J., Vredalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai

,J., Uterback,T., Van Aken,S., Ronning,C.M., Fraser,C.M., Martin

,G.B., Tanksley,S.D. and Giovannoni,J.

Generation of ESTs from tomato callus tissue (2001)

Unpublished (2001)

Contact: CUGI

Clemson University Genomics Institute

100 Jordan Hall, Clemson, SC 29634, USA

Email: <http://www.genome.clemson.edu/orders/index.html>

This clone is available through the Clemson University Genomics

Institute

Seq primer: T3.

Location/Qualifiers

1..661

/organism="Lycopersicon esculentum"

/cultivar="TA496"

/db_xref="taxon:4081"

/clone="CLEC78B18"

/clone_lib="tomato callus"

/tissue_type="callus"

/dev_stage="25-40 days old"


```

SOURCE      potato.
ORGANISM    Solanum tuberosum
REFERENCE   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
AUTHORS     Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
             Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
             1 (bases 1 to 573)
             van der Hoeven,R., Bezzerides,J., Bachem,C., Visser,R., Cho,J.,
             Chiemingo,A., Bougri,O., Buell,C.R., Ronning,C., Tanksley,S. and
             Baker,B.
TITLE       Generation of ESTs from in vitro grown microtubers
JOURNAL     Unpublished (2001)
COMMENT     Contact: Cathy Ronning
             The Institute for Genomic Research
             For clone info: please contact Research Genetics, Libraries
             Division tel 1-800-711-6195, email cdna@resgen.com
             Seq primer: M13F-R.
FEATURES
source      Location/Qualifiers
            1..573
               /organism="Solanum tuberosum"
               /cultivar="Bintje"
               /db_xref="taxon:4113"
               /clone="CSTE19N7"
               /clone_lib="CSTE"
               /issue_type="axillary buds of stem explants; growing
               sink-tubers"
               /dev_stage="7, 8 and 10 days"
               /lab_host="SOLR"
               /note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; Tissue supplied by Christian Bachem and Richard
Visser (Department of Plant Breeding, Wageningen
University, The Netherlands). The cSTA libraries will
attempt to capture the induction and initiation/initial
growth of the tuber in an in vitro system as described in
Bachem et al. (Plant Journal, 1996). Small microtubers
develop from axillary buds attached to stem explants when
placed on a high sucrose medium (10%). Visible
morphological changes occur synchronously at day five in
the axillary buds. The first library, cSTA (1-20) consists
of axillary buds harvested on days 1-3. This targets
those genes involved in induction of the microtubers. The
following libraries, cSTA (21-40) and cSTA (41-60),
capture genes involved in tuber initiation and outgrowth.
This library is noted as P3 in Tanksley lab notebooks."
BASE COUNT  178 a          117 c          99 g          179 t
ORIGIN
alignment_scores:
Quality:      250.00           Length:        185
Ratio:         2.119              Gaps:         6
Percent Similarity: 63.784      Percent Identity: 34.595
alignment_block:
US-09-805-919-1 x BII179655 ..
Align seg 1/1 to: BII179655 from: 1 to: 573
6 LeulepheAlacysValleulleLeuValThrleuleserSerly 22
||| ::| :||||::||| ::| :||||::||| ::| :
29 CTTAAATACCTTACAATTATCCCTTCGAATCATCTTCCGCAACATCCA 78
22 salaginsplesergincysvalProserSerCyeglyAspilegn. 38
:||| ||| :||| :||| ||||:|||||:|||||:
79 TGCTTGAGAGATC...AAGTGCTCGATTTCTGTGGCGATAATCCGTA 125
39 ..llelsphePropheargleuArgThraspprogluhiscyglyarg 54
||||: ||||||:||||:||||:||||:||||:||||:
126 ACATAAGACACCCTTTCATTATAAACACCGATCCAACATGTCATTT 175
55 ArgGLYTYrGIuleuaspCYcylasnasnglnThrValpheasnTyrrLY 71
||| ||||| |||:|||||:|||||:|||||:||||:
176 TCAGGACTGGAATTAGCTGTGAGAAGTAACCAACAGATTATATGTTATC 225

```

```

71  sSerArgIlePheAspValGlnGluIleAsnTyrArgSerTyrSerIleA 88
   ::::: ::::: ::::: ::::: ::::: :::::
226 CTCCAAGAGAGTTGCACCGTTCAAAAGCATCGACTATGATAATCATAAATTC 275
   ::::: ::::: ::::: ::::: ::::: :::::
88  rgleuleuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPhe 104
   ::::: ::::: ::::: ::::: ::::: :::::
276 ACCTGGTAGATCCGACTTTACAAACACAGATGATCATGTCTCTCATATA 325
   ::::: ::::: ::::: ::::: ::::: :::::
105 ProAsn.....HisArgAlaSerTyrAspAlaMetThrSergl 117
   ::::: ::::: ::::: ::::: ::::: :::::
326 CCTTCTAGTTTACCTTTCACAAATACAGTAGTTTCTTCAGCGCTCGTA 375
   ::::: ::::: ::::: ::::: ::::: :::::
117 nIlePheGluTrpValArg...ValAsnAsnAspIleAsnTyrValAsnC 133
   ::::: ::::: ::::: ::::: ::::: :::::
376 CTATTATACATCAGATCGTACTGTACACAGCGCCGATTTTCATGTTTAACT 425
   ::::: ::::: ::::: ::::: ::::: :::::
133 yslLeuAlaProIleAsnSerSerglntYrIleProThrSerPheCysSer 149
   ::::: ::::: ::::: ::::: ::::: :::::
426 GTCCATTGCTGTTAATGATTCACACATTGTGGAATATGAGGCTGCAAA 475
   ::::: ::::: ::::: ::::: ::::: :::::
150 LysAsnSerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAl 165
   ::::: ::::: ::::: ::::: ::::: :::::
476 CTAAGCAGA.....TACACTATTATTAGAGATTGAGAG...ATGGCAGT 516
   ::::: ::::: ::::: ::::: ::::: :::::
166 aSerAspLeuAlaGlyGlyCysArgValGluThrValAlaTrpSerSera 183
   ::::: ::::: ::::: ::::: ::::: :::::
517 CTCTAAGAGTGAGTGATGATGCAGACAGACAGATTATAGGCTTCACCTCAT 566
   ::::: ::::: ::::: ::::: ::::: :::::
183 laPro 184
   :::::
567 GGCTT 571

seq_name: gb_est2:BG455609

seq_documentation_block:
LOCUS      BG455609                655 bp      mRNA      linear      EST 19-MAR-2001
DEFINITION NF061D05PL1F1044 Phosphate starved leaf Medicago truncatula cDNA
ACCESSION  BG455609
VERSION    BG455609
KEYWORDS   BG455609.1  GI:13378934
SOURCE     EST.
ORGANISM   barrel medic.
            Medicago truncatula
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifolieae;
            Medicago.
REFERENCE  1 (bases 1 to 655)
AUTHORS   Liu,J., Scott,A.D., Harris,A.R., Gonzales,R.A., Bell,C.J., Flores
          ,H.R., Imman,J.T., Weller,J.W., May,G.D. and Harrison,M.J.
          Expressed Sequence Tags from the Samuel Roberts Noble Foundation
          Medicago truncatula phosphate-starved leaf library
          Unpublished (2000)
TITLE     Medicago truncatula phosphate-starved leaf library
JOURNAL   Plant Biology Division
          The Samuel Roberts Noble Foundation
          2510 Sam Noble Parkway, Ardmore, OK 73402, USA
          Tel: 580 221 7325
          Fax: 580 221 7380
          Email: mjharrison@noble.org
          Insert Length: 655 Std'Error: 0.00
          Plate: 061 row: D column: 05
          Seq primer: TCACACAGGAACAGCTATGAC.
FEATURES
Source    Location/Qualifiers
            1..655
            /organism="Medicago truncatula"
            /db_xref="taxon:3880"
            /clone="NF061D05PL"
            /clone_lib="Phosphate starved leaf"
            /tissue_type="leaf"
            /dev_stage="trifoliolate"
            /note="Vector: lambda Zap; At the trifoliolate stage, M.
            truncatula plants were transplanted to phosphate-free sand

```


and grown for a further 30 days. During this 30 day period, the plants were fertilized twice weekly with 1/2 Hoaglands solution containing only 20uM potassium phosphate. RNA was prepared from above ground tissues."

BASE COUNT

210 a 141 c 114 g 189 t

1 others

alignment_scores:

Quality: 234.50 Length: 200
Ratio: 2.132 Gaps: 8
Percent Similarity: 55.000 Percent Identity: 33.000

alignment_block:

US-09-805-919-1 x BG455609 ..

Align seg 1/1 to: BG455609 from: 1 to: 655

2 LeuThrArgGlyLeuLeuPheAlaCys..... 10
:::||||| ||::||
20 ATGAGCAGAGGGAATCCTTCACCTGTATGAGGAATCAATATACAGTTT 69
11ValThrLeuIleSerSerSerLysAlaGlnAspIleSerGln 14
70 TAAGCTTCCTCCAATTGTGACNGGTGTACTGTGTACTATTCATC 119
15ValThrLeuIleSerSerSerLysAlaGlnAspIleSerGln 28
120 ATCATGATCAACATGTGATGCAATTACTAATACAAATAACCAACATAT 169
29 CysValProSerSerCysGlyAspIle..GlnIleLysPheProPheAr 44
||| ||||| ||||| ||| ::||:: |||||
170 TGTCCACCTTCTTCATGCGGCAAAATCACAACATATCACCCTTTTCG 219
44 gLeuArgThrAspProGlnHisCysGlyArgArgGlyTyrGlnLeuAsp 61
||| ::||| ||||| ||||| |||||
220 ACTTATGATGACCAACCAACAGCTGTGTATCCGAGTACAGATTATCCT 269
61 YsgLnaAsnGlnThrValPheAsnTyrLysSerArgIlePheAspVal 77
||| ||||| ||||| ::||| ||| ::|||
270 GCGAAACACATCACCGCTGTAACTCTGTTTCAGCTAAATACTATGTG 319
78 GlnGluIleAsnTyrArgSerTyrSerIleArgLeuLeuAspProGlyLe 94
||| ||||| ||||| ::||| ||||| |||||
320 CAGGAATCACTACGTAAATACACAATTCCGCTAGTTCACAGTAAT 369
94 uAsnAspGlnArgGluAsnCysThrValPheProAsnHisArgAlaSer 111
: ::::: ||||| ::||| ::|||
370 T.....GAGAGGGTAAATGCTCCTCCATTCCTCCATATTCTTACTA 413
111 YrAspAlaMetThrSerGlnIlePheGlnTyrValArgValAsnAsp 127
::: ||||| ::||| ::|||
414 AATCCAATTTCACAAGT.....TCTTTCATTCATATCAACAACAGAG 454
127 127
455 GATCCATATCAAAATATGATGTTTGTATTTCTGACCAACCGATACGG 504
128 ...IleAsnTyrValAsnCysLeuAlaProIleAsnSerSer..GlnT 142
||| ||||| ::||| ::|||
505 GCACATATATACTTGAATTGTAGCAACAAGTGAATGATGATCCTGAGT 554
142 YrIleProThrSerPheCysSerLysAsnSerThrGlyPheSerTyrLeu 158
||::: ||::: ||| ||||| ::|||
555 ATGTGATACAGCTCCCTGTATC...AACTCCGACTCCGAAAGCTATCTT 601

seq_name: gb_est1:AW560216

seq_documentation_block:

LOCUS AW560216

547 bp mRNA

linear EST 07-SEP-2000

DEFINITION EST315264 DSIR Medicago truncatula cDNA clone pdsir-26012, mRNA

sequence.

ACCESSION . AW560216

VERSION
KEYWORDS
SOURCE
ORGANISM

AW560216.1 GI:7205642
EST.
barrel medic.
Medicago truncatula
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids I; Fabales; Fabaceae; Papilionoideae; Trifoliaceae;
Medicago.
1 (bases 1 to 547)

REFERENCE
AUTHORS
TITLES

Fedorova,M., Pierson,B.L., Samac,D.A., Vance,C.P., Gantt,G.S., Peng
/H., Ellis,L., Town,C.D., Bowman,C.L., Craven,M.B., Hansen,T.S.,
Holt,I.E. and Fraser,C.M.
ESTs from roots of Medicago truncatula after inoculation with
Phytophthora medicaginis
Unpublished (1999)

JOURNAL
COMMENT

Contact: Carroll P. Vance
Department of Agronomy and Plant Genetics
University of Minnesota
411 Borlaug Hall, 1991 Upper Buford Circle, St. Paul, MN 55108 USA
Tel: 612 625 5715
Fax: 651-649-5058

Email: vance004@maroon.tc.umn.edu
Minnesota EST name:M251571e ; TIGR sequence name:MTBAV90TK ; More
information, including clone ordering, is available at.
'http://chrysis.tamu.edu/medicago/
seq primer: SKmod (CTA gAA CTA gTg gAT CC).

FEATURES
source

Location/Qualifiers

1..547
/organism="Medicago truncatula"
/cultivar="genotype A17"
/db_xref="taxon:3880"
/clone="pdsir-26012"
/clone_lib="DSIR"
/tissue_type="roots infected with Phytophthora
medicaginis"
/dev_stage="roots harvested at 10 days post inoculation
with Phytophthora medicaginis"
/lab_host="E. coli strain XLOLR"
/note="Vector: pBluescript SK-; Site_1: EcoRI; Site_2:
XhoI; cDNA was prepared from polyA+ enriched RNA from
roots harvested at 10 days post inoculation with
Phytophthora medicaginis. The cDNA was directionally
ligated into the Uni-ZAP XR vector from Stratagene and
packaged using GigaPack III Gold packaging extracts.
Plasmids containing cDNA inserts were excised from the
recombinant lambda-ZAP phage using Ex-Assist helper phage
and propagated in XLOLR cells. Note: EST may be of fungal
origin."

BASE COUNT 164 a 122 c 80 g 181 t

ORIGIN

alignment_scores:

Quality: 233.50 Length: 179
Ratio: 2.245 Gaps: 7
Percent Similarity: 58.101 Percent Identity: 33.520

alignment_block:

US-09-805-919-1 x AW560216 ..

Align seg 1/1 to: AW560216 from: 1 to: 547

28 GincysValProSerSerCysGlyAspIle..GlnIleLysPheProph 43
||| ||||| ||||| ||| ::||:: |||||
17 CAATCTTGTCTCTTCCTGCGAAAAAATTGCAACATTAATCAATCCATT 66
43 eArgLeuArgThrAspProGlnHisCysGlyArgArgGlyTyrGluLeu 60
||| ||||| ::||| ||||| |||||
67 CCGACTAAACACAGCAGACACACACTGTGCAACAATTTGTACAGTTGA 116
60 spCysGlnAsnAsnGlnThrValPheAsnTyrLysSerArgIlePheasp 76
::||| ||||| ||||| ::|||
117 ATGCGTAAACACAGCTTACAGTATGAAAGCTGTTGATGCTGAATATTTT 166


```

77 valgingluileasnryargserTyrSerileargleuLeuAspProG1 93
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
167 GTGCAATCAATCAATCAACAATACACAATACAGAGTTGTGACCT... 214
93 yLeuAsnAspGlnArgGluAsnCysThrValPheProAsn..... 106
||| ||| :|||:|||||:|||||:|||||:|||||:|||||
215 ...AATATCAACCAACAATGTCTCCTCCTCCCTCGCTTCTTAT 260
106 ..... 106
261 ACCGACCAATTCACCTATATATACTAATCAATACGAATACAGCTTACT 310
107 HisArgAlaSerTyrAspAlaMet.....ThrSerGlnIlePh 119
::: :|||:|||||:|||||:|||||:|||||:|||||:|||||
311 CGATCAAGTCTTACGATGGAATGATTTATCTTCCACCTCTTATGTTT 360
119 eglutrpValArgValAsnAsnAspIleAsnTyrValAsnCysLeuAlaP 136
|::: :||:|||||:|||||:|||||:|||||:|||||:|||||
361 CGATGGAATTGATTTATCTAGGCCCTGTAATTTACATGAAGTGTACAGTC 410
136 rolleAsnSerSer.....GlnTyrIleProThrSerPheCysSer 149
|| :||:|||||:|||||:|||||:|||||:|||||:|||||
411 CACCTAGTACAAATGTGAATATGATTTATACGATACATCTTCTGCAATG 460
150 LysAsnSerThrGlyPheSerTyrLeuValIleArgGluIleLeuGlnAl 166
::: :||:|||||:|||||:|||||:|||||:|||||:|||||
461 GGTCAACACACAC.....AATATATGCTATTGTTGGGGAC...CCACA 501
166 aserAspLeuAlaGlyGlyCysArgValGluThrVal 178
::: ||| |||||:|||||:|||||:|||||:|||||:|||||
502 TGGATATATTGGAGCCTCAATGTCGTGTTAAGCTTGT 538
seq_name: gb_est1:AW906806
```

```

seq_documentation_block:
LOCUS AW906806 543 bp mRNA linear EST 24-MAY-2000
DEFINITION EST342929 potato stolon, Cornell University Solanum tuberosum cDNA
ACCESSION AW906806
VERSION AW906806
KEYWORDS EST.
SOURCE potato.
ORGANISM Solanum tuberosum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Asteridae; euasterids I; Solanales; Solanaceae; Solanum.
REFERENCE
AUTHORS van der Hoeven,R., Bezzerides,J., Bachem,C., Horvath,B., Visser,R.,
Holt,I.E., Liang,F., Hansen,T.S., Utterback,T., Bowman,C.L., Doan
,B., Bougri,O., Buehl,C.R., Ronning,C.M., Tanksley,S.D. and Baker
,B.
TITLE Generation of ESTs from potato swelling stolons
JOURNAL Unpublished (1999)
COMMENT Contact: Research Genetics, Libraries Division
Tel: 1-800-711-6195
Email: cdna@resgen.com
5 prime sequence.
FEATURES
source
1..543
location/Qualifiers
/organism="Solanum tuberosum"
/cultivar="Bintje"
/db_xref="taxon:4113"
/clone_lib="cSTRA4F9"
/clone_lib="potato stolon, Cornell University"
/tissue_type="axillary buds of stem explants, swelling
stolons"
/dev_stage="1 to 3 days"
/lab_host="SOLR"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; RNA was supplied by Christian Bachem & Beatrix
Horvath(Laboratory of Plant Breeding, Dept. of Plant
Sciences, Wageningen University, The Netherlands). Total
```

RNA was isolated from developing axillary buds of potato nodal stem cuttings cultured on medium for the introduction of tuber formation as described in Bachem et al. (Plant Journal 1996). Tissue samples were taken of stages corresponding to growing stolons and the early stages of tuber formation."

```

alignment_scores:
Quality: 233.00 Length: 158
Ratio: 2.284 Gaps: 3
Percent Similarity: 64.557 Percent Identity: 32.911
alignment_block:
US-09-805-919-1 x AW906806 ..
Align seg 1/1 to: AW906806 from: 1 to: 543
```

```

7 leupheAlAcysValleuLeuValThrIleSerSerSerLysAl 23
|||||:|||||:|||||:|||||:|||||:|||||:|||||
23 TTGCTTCTTACATCTCATATTCATGTACCTCTCAGAAACCCATTAGC 72
23 aglAspIleSerGlnCysValProSerSerCysGlyAspIle...GlnI 39
|::: :||:|||||:|||||:|||||:|||||:|||||:|||||
73 AGACAAAGACAGACAGTGTGTTCTCTTCAAGTTGTGGCATATTCATAACA 122
39 lelySpheProPheArgLeuArgThrAspProGluHisCysGlyArgArg 55
|::: :||:|||||:|||||:|||||:|||||:|||||:|||||
123 TCAGTTATCCCTTCGATTTGAAAGTGAATCCAAAGCAGCTGTGGCCTTAA 172
56 GlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPheAsn..... 69
::: ||| |||||:|||||:|||||:|||||:|||||:|||||
173 AATTATGAATTGAGCTGTGAAGAGATCACACCATATTCACATACTACCTCC 222
70 .....TyrLysSerArgIlePheAspValGlnGluIleAsnTyrA 83
||| ||| :||:|||||:|||||:|||||:|||||:|||||
223 ACATGATTGGTATGATTCACCTGAAGTACTATGTTCACAGCCATCAATTACG 272
83 rgseryTyrSerIleArgLeuLeuAspProGlyLeuAsnAspGln..... 97
::: :||:|||||:|||||:|||||:|||||:|||||:|||||
273 ATAACCTCAGTATCCGCTGTTGTGATCCCTGTTAAGAGAACAGTTATT 322
98 .....ArgGluAsnCysThrValPheProAsnHisArgAlaSerTy 111
::: ||| :||:|||||:|||||:|||||:|||||:|||||
323 TGCTCTCTTCTCAGCAATTCGATTACGATTACACTACCCCTCTTGCTT 372
111 rAspAlaMetThrSerGlnIlePheGluThrValArgValAsnAsnAspI 128
|::: :||:|||||:|||||:|||||:|||||:|||||:|||||
373 TGATTTCTTCTATATATACAGGCCCTGATCAGGTTATGTAAACAGTACCAA 422
128 leasnTyrValAsnCysLeuAlaProIleAsnSerSerGlnTyrIlePro 144
|::: :||:|||||:|||||:|||||:|||||:|||||:|||||
423 TCAGTCTCTTCAGCTGTCATTGTCCTTAATTCTCCAGTCTTGTGGA 472
145 ThrSerPheCysSerLysAsnSer 152
::: ||| :||:|||||:|||||:|||||:|||||:|||||
473 ATCACTAATTGTTCAACAGGAGT 496
```

```

seq_name: gb_est2:BI922909
seq_documentation_block:
LOCUS BI922909 710 bp mRNA linear EST 18-OCT-2001
DEFINITION EST542813 tomato callus Lycopersicon esculentum cDNA clone cLEC78N4
ACCESSION BI922909
VERSION BI922909.1 GI:16221166
KEYWORDS EST.
SOURCE tomato.
ORGANISM Lycopersicon esculentum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
```

REFERENCE 1 (bases 1 to 710)
Alcala,J., Vrebalov,J., White,R., Vision,T., Karamycheva,S.A., Tsai
J., Utterback,T., Van Aken,S., Ronning,C.M., Fraser,C.M., Martin
G.B., Tanksley,S.D. and Giovannoni,J.
Generation of ESTs from tomato callus tissue (2001)
Unpublished (2001)
COMMENT
JOURNAL
TITLE
Clemson University Genomics Institute
Clemson University
100 Jordan Hall, Clemson, SC 29634, USA
Email: <http://www.genome.clemson.edu/orders/index.html>
This clone is available through the Clemson University Genomics
Institute
Seq primer: T3.

FEATURES
source
Location/Qualifiers
1..710
/organism="Lycopersicon esculentum"
/cultivar="TA496"
/db_xref="taxon:4081"
/clone="CLEC78N4"
/clone_lib="tomato callus"
/tissue_type="callus"
/dev_stage="25-40 days old"
/lab_host="XLI-Blue MRF"
/note="Vector: pBluescript SK(-); Site_1: EcoRI; Site_2:
XhoI; supplier: Boyce Thompson Institute; sequencing: The
Institute for Genomic Research; CLEC - Cotyledons of
seedlings 7-10 days post-germination were excised, cut at
both ends and placed on MS medium with no selection. Mixed
callus was harvested at 25 and 40 days and included
undifferentiated masses. Tomato Callus EST Library"

BASE COUNT 204 a 158 c 131 g 217 t
ORIGIN

alignment_scores:
Quality: 232.00 Length: 165
Ratio: 2.071 Gaps: 6
Percent Similarity: 67.879 Percent Identity: 35.152

alignment_block:
US-09-805-919-1 x BI922909 ..

Align seg 1/1 to: BI922909 from: 1 to: 710

```
31 ProSerSerCysGlyAspIleGln..IleLysPheProPheArgLeuAr 46
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||:
108 CCTTCGTGCTGTGTCATATCCGAGACATCAGCTACCCCTTTTCACCTTAA 157
46 gThrAspProGluHisCysGly..ArgArgGlyTyrGluLeuAspCysG 62
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
158 CAGTGACCCAGAAATTTGTGGAGATGATCCGAAATTTGAATTAGTTGTG 207
62 InAsnAsnGlnThrValPheAsnTyrLysSerArgIlePheAspValGln 78
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
208 AAGATAACCAACGCTTATGTGCATCCTCTCCAGAGAGCTGTATGTGCAA 257
79 GluIleAsnTyrArgSerTyrSerIleArgLeuLeuAspProGlyLeuAs 95
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
258 GCCATCAACTATAATAGTAGACAATTCACCTGTAGATCCAGCTTTACA 307
95 nAspGlnArgGluAsnCysThrValPheProAsn.....HisArgAlas 110
:||||:||||:||||:||||:||||:||||:||||:||||:
308 AACACAGAGATGATCTATGCTTTCAAGTCCCTCAGCTTCTGTCTTCGACC 357
110 erTyrAspAlaMetThrSerGlnIlePheGluTyrPValArgValAsnAsn 126
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
358 AATCCGATACAACTCTCCGATCATACTATAGTGGCTTAGATCAGCAGAG 407
127 AspIleAsnTyrValAsnCysLeuAlaProIleAsnSerSerGln..Ty 142
|||||:|||||:|||||:|||||:|||||:|||||:|||||:
```

```
408 CCCATTTTCATGTCACACTGTCCATCTGCTGTTAATAGTTCTTCGACATT 457
142 rIleProThrSerPheCysSerLysAsnSerThrGlyPheSerTyrLeuV 159
:::|||||:|||||:|||||:|||||:|||||:|||||:|||||:
458 TCTGAAATAGCTGGCTGCAGAAATTAAGCAGG.....TACACTTATTAA 501
159 aIleArgGluIleLeuGlnAlaSerAspLeuAlaGlyIcysArgVal 175
||| ||| :|||:|||||:|||||:|||||:|||||:|||||:
502 AGATTGAGAA...ACGAAAGTCTCTGAAGTCAGCCATGGATGACAGAAATG 548
176 GluThrValAlaTyrPheSerAlaProGlyIleSerSerAsnLys 190
||| :|||:|||||:|||||:|||||:|||||:|||||:|||||:
549 GAATTTATAGCATGACCTCATTTGCTGATATTAACGACGAGAGA 593
```

OM of: US-09-805-919-1 to: GenEmbl:* out_format : pfs
Date: Jun 4, 2002 11:16 PM

About: Results were produced by the Gencore software, version 4.5,
Copyright (c) 1993-2000 Compugen Ltd.

Command line parameters:

-MODE=frame+pn.model -DEV=xlh
-Q=/cgcn2_1/USPTO.spool/US09805919/runat_04062002_121211_7635/app_query.fasta_1.369
-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -GAPOP=12.000 -GAPEXT=4.000
-MINMATCH=0.100 -LOOPEXT=0.000 -LOOPEXT=0.000 -OGAPOP=4.500
-OGAPEXT=0.050 -XGAPOP=10.000 -XGAPEXT=0.500 -FGAPOP=6.000
-FGAPEXT=7.000 -YGAPOP=10.000 -YGAPEXT=0.500 -DELOP=6.000
-DELEXT=7.000 -START=1 -MATRIX=blotsum62 -TRANS=human40.cdi
-LIST=45 -DOCCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0
-ALIGN=15 -MODE=LOCAL -OUTFMT=pfs -NORM=ext -HEADSIZE=500
-MINLEN=0 -MAXLEN=200000000 -USER=US09805919 -CGCN1_1_9453
-NCPU=6 -ICPU=3 -LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30
-NO_XLPHY -WAIT -THREADS=1

Search information block:

Query: US-09-805-919-1
Query length: 308
Database: GenEmbl:*
Database sequences: 1797656
Database length: 187333701
Search time (sec): 1845.600000

Sequence	Strd	Orig	Zscore	EScore	Len	Documentation
gb_pl:AF077130	-	195.00	324.68	1.0e-09	4392	AF077130 Oryza sativa receptor
gb_pl:AF238476	+	193.00	323.03	1.2e-09	3628	AF238476 Oryza sativa receptor
gb_pl:AP003338	-	193.00	286.08	1.4e-07	112222	AP003338 Oryza sativa genom
gb_pl:AP003219	+	190.00	279.24	3.4e-07	126434	AP003219 Oryza sativa genom
gb_pl:AP003219	-	188.50	276.47	4.9e-07	126434	AP003219 Oryza sativa genom
gb_pl:AP003250	-	174.00	248.03	1.9e-05	146670	AP003250 Oryza sativa genom
gb_pl:AP002747	-	167.50	235.91	8.9e-05	147857	AP002747 Oryza sativa genom
gb_pl:AF100765	+	166.00	274.86	6.0e-07	3067	AF100765 Oryza sativa receptor
gb_pl:AF238475	+	163.50	267.37	1.6e-06	3999	AF238475 Oryza sativa receptor
gb_htg:AC091086	-	163.50	228.69	0.0002	145358	AC091086 Oryza sativa chrom
gb_pl:AC006418	-	163.00	235.98	8.8e-05	67730	AC006418 Arabidopsis thaliana
gb_pl:AF237569	+	162.00	266.67	1.7e-06	3300	AF237569 Oryza sativa receptor
gb_pl:AP002867	-	162.00	225.99	0.0003	144322	AP002867 Oryza sativa genom
gb_pl:AF238473	+	161.50	259.30	4.4e-06	6000	AF238473 Oryza sativa receptor
gb_pl:AP004493	+	161.50	239.98	5.3e-05	36100	AP004493 Lotus japonicus genc
gb_pl:AP002541	-	160.50	223.12	0.0005	145576	AP002541 Oryza sativa genom
gb_pl:AP003450	+	159.50	219.58	0.0007	170226	AP003450 Oryza sativa genom
gb_pl:AF237569	+	157.00	225.99	0.0003	144322	AF237569 Oryza sativa receptor
gb_pl:AP003450	-	157.00	214.95	0.0013	170226	AP003450 Oryza sativa genom
gb_pl:AC006300	+	153.50	215.78	0.0012	86348	AC006300 Arabidopsis thaliana
gb_pl:AP003250	+	152.50	208.23	0.0031	146670	AP003250 Oryza sativa genom
gb_pl:AB025622	-	149.50	211.87	0.0019	62420	AB025622 Arabidopsis thaliana
gb_pl:AF235196	-	147.00	213.21	0.0016	35872	AF235196 Triticum aestivum PS
gb_pl:AC083891	-	146.00	211.59	0.0020	35116	AC083891 Arabidopsis thaliana
gb_pl:AC013288	-	146.00	200.66	0.0082	96899	AC013288 Arabidopsis thaliana
gb_pl:TAU78762	+	145.50	238.11	6.7e-05	2744	U78762 Triticum aestivum recep
gb_pl:AF085167	+	143.50	234.00	0.0001	2849	AF085167 Hordeum vulgare recep
gb_pl:AC013354	-	142.50	194.36	0.0184	95327	AC013354 Genomic sequence for
gb_pl:AC079281	+	141.50	191.72	0.0257	102540	AC079281 Arabidopsis thalian
gb_pl:AF237570	+	139.00	222.02	0.0005	3999	AF237570 Oryza sativa receptor
gb_pl:AC069471	-	138.50	183.86	0.0705	126991	AC069471 Arabidopsis thalian
gb_pl:AF237567	+	137.00	219.01	0.0008	3751	AF237567 Oryza sativa receptor
gb_pl:AB005247	+	136.00	183.20	0.0767	87841	AB005247 Arabidopsis thaliana
gb_pl:FL1F1	-	136.00	179.55	0.1226	123386	AC002131 Arabidopsis thalian
gb_pl:AF237546	+	135.00	230.97	0.0002	876	AF237546 Avena sativa receptor
gb_pl:AF164021	-	135.00	211.97	0.0019	5112	AF164021 Oryza sativa clone GR
gb_pl:AP002747	+	135.00	175.75	0.1996	147857	AP002747 Oryza sativa genom
gb_pl:AF237548	+	132.00	225.39	0.0003	878	AF237548 Avena sativa receptor
gb_pl:AY052245	+	132.00	215.73	0.0012	2153	AY052245 Arabidopsis thal
gb_pl:AP002541	+	132.00	170.36	0.3983	145576	AP002541 Oryza sativa genom

gb_pl:AF085166	+	131.50	212.95	0.0017	2559	AF085166 Hordeum vulgare
gb_pl:AF100769	+	131.50	212.18	0.0019	2748	AF100769 Zea mays receptor
gb_pl:AF325197	+	131.50	190.41	0.0304	20754	AF325197 Triticum aestiv
gb_pl:AY051085	+	131.00	221.91	0.0005	1021	AY051085 Arabidopsis thal
gb_pl:AF360349	+	131.00	220.00	0.0007	1220	AF360349 Arabidopsis thal

seq_name: gb_pl:AF077130

seq_documentation_block:

LOCUS AF077130 4392 bp DNA linear PLN 25-JUL-1998
DEFINITION Oryza sativa receptor-like protein kinase gene, complete cds.

```

1145 ACAACCACTTCGCTTCCTTCCTGCGACACACTTGAAGTATCAGCT 1096
      : ||||| ||||| : : : : || : : : :
41 heProPheArGLeuArGthrAspProGluHisCysGlyArGArgGlyTyr 57
      : ||||| ||||| : ||||| ||||| : : : : ||
1095 ACCCTTCCGCTTACAAAGGTATCCACTTGGCTGGTGGTATGAAGCTTAT 1046
      : ||||| ||||| : : : : : : : : : : : : : :
58 GluLeuAspCysGluAsnAsnGlnThrValPheAsnTyrLysSerArgI 74
      : ||||| ||||| : : : : : : : : : : : : : :
1045 GAGCTAGTGTGTAGAGATGGCAGAGCTATATTCACATCAACACAGAGAA 996
      : : || : : || : : || : : || : : || : : ||
74 epheAspValGlnGluLeuAsnTyrArGSerTyrSerIleArGLeuA 91
      : : || : : || : : || : : || : : || : : ||
995 ATATTGTGACAGACATCTCCATATAAGAGTCAATTTGGGTGTG 946
      || : : || : : || : : || : : || : : ||
945 ATGCTAATTAGACAC...AGCAGTGGCCCTCTCCACTGTGGAAT 902
      : : : : : : : : : : : : : : : : : :
107 HisArgAlaSerTyrAspAlaMetThrSerGlnIlePhe..... 119
      : : : : : : : : : : || : : : : : : : :
901 AATCTTCCCTACTTCAATGACATGGGCACCTAAGCTGTACACTTCTGCAGT 852
      : : || : : || : : || : : || : : || : : ||
120 .GluTrpValArGValAsnAsnSpIleAsnTyrValAsnCysLeuAlaP 136
      : : || : : || : : || : : || : : || : : ||
851 TCGGTGGCT.....ACATTCTTAATTGCTCACGGG 820
      : : || : : || : : || : : || : : || : : ||
136 roIleAsnSerSerGlnTyrIleProThrSerPheCysSerLysAsnSer 152
      : ||||| : : : : : : : : : : : : : : || : : : : : :
819 CGATAACAATGGCATGTACATGCCCTGTGCT..TGCTTGAGTGGGAAT 773
      : : || : : || : : || : : || : : || : : ||
153 ThrGlyPheSerTyr..LeuValIleArGLeuGluLeuGlnAlaSerAs 168
      : ||||| ||||| ||||| : : : : : : : : : : : :
772 ACTTCTTTGTTTATGCTCTGACTACATCGAGCTTACTATGTTCAAA 723
      : : || : : || : : || : : || : : || : : ||
168 pLeuAlaGlyGlyCysArGValGluThrValAlaIlePheSerSerAlaProG 185
      : : || : : || : : || : : || : : || : : ||
722 CATCGAGCTTCTGTGATACTGGCTGTG.....ATTCCTG 685
      : : || : : || : : || : : || : : || : : ||
185 IyIleSerSerAsnLysSer.....SerThrIleSerSerThr 197
      : : : : : : : : : : : : : : : : : : : :
684 TGGATGATCGTACAAAGAATGTACCAGACTATGCAAGCTATGCAGATGTT 635
      : : || : : || : : || : : || : : || : : ||
198 HisGlnGlyLeuAlaTyrGlyPheGluLeuSerTrpLysArgAsn.... 212
      : : : : : : ||||| : : || : : || : : ||
634 GTGAATTCATGAGGAATGTTTTCCTGTTCTGTTCTCGAGCTTGCTT 585
      : : : : : : ||||| : : || : : || : : ||
213 .....LeuIleuCysArgAsnCysA 219
      : : : : : : : : : : : : : : : : : : : :
584 GTTTGATCATCTTATAAAAAATATATATTCCTCTTTGCAATTCTGCAG 535
      : : || : : || : : || : : || : : || : : ||
219 spArGSerArGlyGlyGlyCysThrIleGlu.GluAsnSerAspArgAl 235
      : : || : : || : : || : : || : : || : : ||
534 AAACCTTAAGG..AGCAATATr..ATAGTAGGAACATCAAAACTGGA 491
      : : || : : || : : || : : || : : || : : ||
235 aThrCysArGTYrTrpCysLysGluAspIleHisValSerLysLeuThrP 252
      : : || : : || : : || : : || : : || : : ||
490 CTCTGCGCAATTATTGGACCGAGCTACAAT.....T 459
      : : || : : || : : || : : || : : || : : ||
252 heArG.CysLysValGluTyrTyrSer.....Va 261
      : : || : : || : : || : : || : : || : : ||
458 TTAGGATGTGAATTAATTATTACTCTAGTGCCACGCAAGTTTGGGTTAC 409
      : : || : : || : : || : : || : : || : : ||
261 lTyrValLeuPhePheGlyGlyIleGlyIleGlyValLeuAla***A 278
      : : || : : || : : || : : || : : || : : ||
408 TCTCGTTCGTGTAATT.....GCCGTAGAGATTGTTA 377
      : : || : : || : : || : : || : : || : : ||
278 rGpHeLeuLeuGlyIleProIleLeuIle 287
      : : : : : : ||||| : : : : : :
376 AGTGCATTAATAGGTACTCCAGATATGCTT 348
      : : || : : || : : || : : || : : || : : ||
seq_name: gb_pl:AF238476

```

```

seq_documentation_block:
LOCUS      AF238476                3628 bp    DNA        linear    PLN 21-JUN-2000
DEFINITION Oryza sativa receptor-like kinase (RGL1) gene, complete cds.
ACCESSION  AF238476
VERSION    AF238476.1    GI:8575485
KEYWORDS
SOURCE
ORGANISM   Oryza sativa.
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 3628)
            Yun,C.-H., Lee,G.-R. and Kim,H.-I.
            Molecular analysis of receptor-like kinase in rice
            Unpublished
            2 (bases 1 to 3628)
            Yun,C.-H., Lee,G.-R. and Kim,H.-I.
            Direct Submission
            Submitted (24-FEB-2000) CytoGenetics, National Institute of
            Agricultural and Technology, Seodundong 249, Suwon 441-707, Korea
FEATURES
            source
                1..3628
                /organism="Oryza sativa"
                /variety="Nipponbare"
                /db_xref="taxon:4530"
                /chromosome="1"
                join(<1327..2066,2348..>3623)
                /gene="RGL1"
                /product="receptor-like kinase"
                <1327..>3623
                /gene="RGL1"
                join(1327..2066,2348..3623)
                /gene="RGL1"
                /note="protein kinase"
                /codon_start=1
                /product="receptor-like kinase"
                /protein_id="AAF78020.1"
                /db_xref="GI:8575486"
                /translation="MHPLSTMQAFVFIYAVLALLAGDVEGRGQYPFSCGHLQNI
                PFRRGDPQRCGVPSYELDCRDSKATIRINTGTYVTSINTYTSFVWVYDASLKDNTS
                SCPLPRSDQLPFVSGGIQSGHGDWLDLDPGATWVSFVNCQAVRNNSVYVPDCL
                STSSFEVYVFGSWMIMPPSVFPIIGNLETSRYLAMIPLGWDSP.LPHNASFSDIVRSM
                RNGFAVHFPIIHRMSRIGHIKDCMGSIILMRIHCTICFCFYDCWFYDSVFLII
                QYITITAPFOCLKTHLKVITVYVKKLFYAVLCRFVIAPLTLLTFLAFKYWKTRIKID
                AVEKFLQMDMLGPTRYAYTDIIAMTSFRDKLGQGVSVFGVYLLPGDVHVAIKML
                SNYNNGEEFISEVSTIGSIHHVNVRLVGYCAEEMRSALVIEYMHGSLDRFIFSPD
                KSLWDKLEIALGIARGINYLHOGCDMLHFDIKPHNILLDSNEVPEKVADEGLAKL
                YPRDNSFMPYSARGTGVYIAPEMISRFEGIISSKSDVYSFGMLLEAGGRNSKQN
                MSSSSQSYPSWYNOLVOQKGEIANFNMHELEKLCVVGILHCTIOMKSHDRPTMSE
                VIEMLEGDVGGLQLPSRPFCDDEPLPLVYDSRFSSELTEISEDE"
BASE COUNT      921 a      738 c      889 g      1080 t
ORIGIN
alignment_scores:
            Quality: 193.00      Length: 243
            Ratio: 1.451      Gaps: 10
            Percent Similarity: 54.733      Percent Identity: 25.926
alignment_block:
US-09-805-919-1 x AF238476 ..
Align seg 1/1 to: AF238476 from: 1 to: 3628
1 LeuphealacysValleuleuValThrleuileSerSerLysal 23
:::|||||::: |||||::: ::::: :::::
:::|||||::: |||||::: ::::: :::::
1366 ATCTTGTCTGTCTTGTCTTGTGACGAGATGTCGAGGGGCGCAAGG 1415
23 aglnaspileserglncysvalprosersercysglyaspilegln..I 39
: ||| ||| ||||| ||| ::::|
1416 G.....TGCATCCCTTCTCCTGTGACATCTCCAGAACA 1450
39 lelysphepropheargleuargThraspProgluHiscysglyargarg 55

```



```
||||: ||||| ||| |||||:|||||
1451 TTCTCATCCTTCCGCCGCGAGGTGATCCACAAGATGCGGTCCG 1500
56 GlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPheAsnTyrLysSe 72
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1501 TCATATGAGTTAGACTGCAGAGATAGCAAGGCTACCATTAGGATCAACAC 1550
72 rArgIlePheaspValGlnGluLeuAsnTyrArgSerTyrSerIleArgL 89
: ||| ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1551 AGGACATACATATGTAACACAGCATCACTACTACTTCTCTCTGGG 1600
89 euleuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPhePro 105
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1601 TTGTGATGCCAGCCTAAAGATACAACACAGCAGCTGC.....CCT 1641
106 AsnHisArgAlaSerTyrAspAlaMetThrSerGlnIlePheGlu.... 120
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1642 CTTCCTCGCTCGGATCAGCTTCTTCTTCGTTCCGGGGGTATCCAGGATC 1691
121 ..... 122 TrpV 122
1692 ACATGGCGGCTGGGACCTGGCCCTTGACCTTGACCTGGAGCTACATGGG 1741
122 aLArgValAsnAsnAspIleAsnTyrValAsnCysLeuAlaProIle.. 137
||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1742 TT.....AGCTTTGTGAATTGTTTCACAGGCAGTAAGG 1773
138 AsnSerSerGlnTyrIleProThrSerPheCysSerLysAsnSerThrG1 154
|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1774 AATAATAGTGTGTACGTTCCCTGTGTGAT..TGCCTGAGCACAAAGCTTTC 1820
154 yPheSerTyrLeuValIleArgGluIleLeuGlnAlaSer..... 167
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1821 GTTGTATTATGTGTTGTTGTTCTCGATATATGCCCTCATCTGTTTCCAA 1870
168 .....AspLeuAlaGlyGlyCysArg.....ValGluThr 177
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1871 TCATCGGAACCTTGAGACTTCATGTCGCTTACCTTGCCATGATTCCTTG 1920
178 ValAlaTrpSerSerAlaProGlyIleSerSerAsnLysSerSerThrLe 194
:|||||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
1921 GGTGTTGGGACTCGCCATTGCCACATATGCAAGCTTTTCAGACATCGT 1970
194 uSerSerThrHisGlnGlyLeuAlaTyrGlyPheGluLeu.....SerT 209
: ||| :||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
1971 CAGATCCATGAGGAATGATTTGCTGTTTCATTTTCCCATATATACATAGT 2020
209 rPlyAsrAsnLeuLeuCysArgAsnCys 218
|||:|||||:|||||:|||||:|||||:|||||:|||||:|||||
2021 GGAGTAGGATGGGCACATCAAGATGTG 2049
```

```
seq_name: gb_pl:AP003338
seq_documentation_block:
LOCUS AP003338 112222 bp DNA linear PLN 13-APR-2001
DEFINITION Oryza sativa genomic DNA, chromosome 1, BAC clone:OU1212_B09.
ACCESSION AP003338
VERSION AP003338.2 GI:13366212
KEYWORDS
```

ORGANISM Oryza sativa (cultivar: Nipponbare) DNA, clone:OU1212_B09.

REFERENCE 1 (sites) Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae; Ehrhartoideae; Oryzeae; Oryza.

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC clone:OU1212_B09

JOURNAL Published Only in DataBase (2001) In press

REFERENCE 2 (bases 1 to 112222)

AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.

TITLE Direct Submission

JOURNAL Submitted (28-FEB-2001) Takuji Sasaki, National Institute of

COMMENT

Agrobiological Resources, Rice Genome Research Program, Kannondai 2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/, Tel:81-298-38-7441, Fax:81-298-38-7468)
On Mar 16, 2001 this sequence version replaced gi:13359074.
The nucleotide sequence of this BAC clone was generated by combining Monsanto and RGP-Japan sequencing data.
Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at RGP. Protein homologies of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein.
The orientation of the sequence is from -21M13 to M13rev of the BAC clone. Detailed information on overlap and assembly quality together with annotation of this entry is available at
http://rgp.dna.affrc.go.jp/Genomeseq.html.

FEATURES
source

```
gene
CDS
        /gene="OJ1212_B09.1"
        /complement(join(575..1850,2132..2871))
        /gene="OJ1212_B09.1"
        /codon_start=1
        /product="receptor-like kinase"
        /protein_id="BAB39434.1"
        /db_xref="GI:13366213"
        /translation="MHPLSTTQQAFFIFAVLALLAGDVEGRGQCYPEPCGHLONISH
        PFRRRBDPQRCGVPSYELDQRDSCATIRINTGTYVTSINYTTSVFVYVDASLKDNTS
        SCPLPSDQLPEVSGGIQSGHGWDLALDPGPAFWVSFVNCQAVERNVSYPVDCI
        STSSSFYVFGSWIMPPSVFPIIGNLETSCRYLAMIPLGWDSP.LPHNASFSDIVRSM
        RNGFAVHFPIIHRWSRIGHIKDCLMGISIRILWRHICTICFCFQVDCWYFYDSVFLII
        QYITTAFFQCLKTHLKVITTVYVKLFYFVAVLCRFYIAPLTLFLFAFKYWKTRIKID
        AVEKFLQMLGLPTRYAYTDIIAMTSHFRDKGGGYSVFKEGVLPGDVHVAIKML
        SNYNGEELFISEVSTIGSIHVNVRVLYGCAEMRSALVYEFMHGSLDRFIFSPD
        KLSWDKLNELIALGIARGINLHQGDMQILHFDIKPHNILLDSNFYKVAADFGLAKL
        YPRDNTFVPSAARGTVGTAPEMISRSFGIISKSDVYSFGMLLLEMAGRRNSKON
        MSSSSQSYVPSWVYNQIVQKMEIYANAFNMHELEKLCVVGILHICQMKSHDRPTMSE
        VIEMLEGDVGGQLPSRPFCDDEPLPLLVDSGRSSELTETISEDE"
        /join(3276..4074,4801..4824,5194..6263)
        /gene="OJ1212_B09.2"
        /join(3276..4074,4801..4824,5194..6263)
        /gene="OJ1212_B09.2"
        /codon_start=1
        /product="putative receptor kinase"
        /protein_id="BAB39435.1"
        /db_xref="GI:13366214"
```

```
/translation="MGTILATAFLLSVLSSHGSYAMASSWDDODFFKHCPPSRCSKH
/PEIRFPRLQSSNTPSSCGSSHAKLICSQGDITLIHHPFLGPKVATLADYKKAVMKII
PGGSSSPCLLHKFNSTNLADVNOLYLTEGRIRVSCSEFTTSRAMIDGXNTV
IADKYVRLIPCLRDTSHFSYLVSTRLLYALPLDCMVVSKGNIPINPYTAGLTFKO
MAERILNSAEITLILLGSIIPYNCRCQOGORCAFSORNOTFCMHGSRVKVIAAT
SVAEYAVSLVAVTLVLSLKQRYNEVHLKVENFLRTGTSKPTRYTFQVKKITRR
EKEVGGGEGFTVYKGKLLNGVPAVKMLNPTGDEDEITTEVATIGRIHANIIHLL
GFCSEGTFRALIVYEFMPNESLEKYLFLHDHNTPOELSLPNKMLDIALGIARMEYLHO
GCNQRILHFDIKPHNILLDYNSPKISDFGLAKLCPRDQSIYTMKARGTMGIAPEL
YSRNGEISYKSDVYSFGMLVLEMVSGRSWPSIKNQNEVYFPEWYIEKVTGOEFV
LSREMTBEKQMVRLALVALMCIQWNPNNRPSMTKYVNMITGRLONIQVPPKPFVSY
```


misc_feature complement(9881..12639)
/gene="OJ1212_B09.3"
/note="probably inactive due to frameshift(s) in CDS
probably inactive due to stop codon(s) in CDS
pseudogene, receptor-like kinase"
/pseudo
complement(9881..12639)
/gene="OJ1212_B09.3"
complement(14471..19744)
/gene="OJ1212_B09.4"
complement(14471..19744)
/gene="OJ1212_B09.4"
/note="probably inactive due to stop codon(s) in CDS
pseudogene, polypeptide"
/pseudo
join(22129..22261,22601..22738,22822..22899,22973..23111,
23667..23800,24540..24922)
/gene="OJ1212_B09.5"
join(22129..22261,22601..22738,22822..22899,22973..23111,
23667..23800,24540..24922)
/gene="OJ1212_B09.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB39436.1"
/db_xref="GI:13366215"
/translation="MHARCGPAMHARRSPMVGSAVETDILQNDSTHKLKAMRKE
AYBERGLQWNRAGRGCSRRASRQWDEDTESFTKRSKHEPNYGQEEKERRR
KKGIWEDDMTQHWMPROCHVGRAPSGHRHVSSETTLQDRKRYKLRENSLGESSYDIY
FKGNIYKDRKVAFOYNTLMHCLFGLISEGLQSCACRADDGEMRTKGYLLSR
DGLITIGSSMSDYLMQHTPOKMQCPNNTGQFSHTITTSQSLYDKRKLPHVHG
KEITRSVKDSSCLPLSALTSGVHSTRVAVIRMPQQRREIRNGLSSIFFOQPSSTF
"
gene complement(join(25076..26181,26293..26423,26594..27309))
/gene="OJ1212_B09.6"
complement(join(25076..26181,26293..26423,26594..27309))
/gene="OJ1212_B09.6"
/codon_start=1
/product="receptor-like kinase"
/protein_id="BAB39437.1"
/db_xref="GI:13366216"
/translation="MRVIVLCVLGVLPDAAGRRHRRHDCPFTGHLSDVSEFP
RRGDPPECVQSELTLCADKATIQIDKETYSVDINYGDSLWVVDASFLDSRSC
LPRWNPPLRDPLOAKSHHIEIAPPGVTWASFVNCSEIRNSSWMPVACLSIS
RSFVYVFTGQSAIQNLSPSCGYLATPILGSKLNSTALQNVSYQDVKLMATGFA
VRPEFTVSGWNPKECLALISRIANAIIDFYFWSCFLGDRSHNLITYMVVDTALL
ILKWTAVLCRFVAPLAVFIFLAHKYWRNKITIDAVEKFLQMLTLPTRYATDLTA
ITGHFGEKLGQGYSVYKGLPGYVAVKYLENANCNGEETSEVSTIGRIHVN
VRLVGFCEELRALVYEMPRGSLDKITFSKRSFSWMDKLNELALGIARGINYLHOG
CDMQILHFDIKPHNILLDNEFVPAVADFLAKLYPRDNSFVNALRGITIGYIAPEMI
SRSGVSSKSDVYSEGMILLLEMAGRRNSDMHAGNSSQAYPSPVYDRLIEQVGVG
EISATVAMHLEKRLCIGLHCIOKSHDRPTMSEVIEMLEGCVGLQMPRPFFC
DDESMSPMDSYQSSGLTEILEDE"
gene join(28136..28907,29728..29817,30034..30057,30164..31236)
/gene="OJ1212_B09.7"
join(28136..28907,29728..29817,30034..30057,30164..31236)
/gene="OJ1212_B09.7"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB39438.1"
/db_xref="GI:13366217"
/translation="MDFNLILTVLLSPLKYESDVATASDDEFEKTCSSQRCSEK
GPEIRYPERLSTQPPSCGAPQMSQSGDTILDHVLGSCKVTMIYRHVIMNAIPL
VDSLPHCPLOKLVSNQSTAVYKPHPTSEVASVYCSRDSIDTNGYIGVPTSCPSLAN
NASQFWLAYPYTYMSILPLGCTIVSKDIPMPYSYDKNGPNFIDISIFETAKRVISG
ETVFTWTSNVTSGCOCCEREGRCGEGSQRNAFCQHHKTRPYPSYTLVRKLCICI
FPAMLCKIESGPRVTIATSSVGTIVLSLIVATALYISLSRYNEIHLKVEMTL
KTYGTSKTRTYTSEVKKIARFKDKLHGAFGTIVKGLPNGVPAVKMLNSVSEG
QEFINEVATIGRIHNAVIRLVGLFCSEGRALYIELMPNESLEKYIFPHGSNISREL
LYPDKMLDIALGIARMEYLHOGCNRILAFDIPHNILLDYSPNPKISDFGLAKLCA
RDOISIVTLTAAGTMGYIAPELYSRNFGLISYKSDVYSGMLYLEMVSGRNTDPTVE
SONEFYFPEWIERVINGODLVLTMETTQGEKEMVRQLAIVALWCIOQWPKDRPSMTK
VVMMLTGLQNLQVPRKPFISSQNLVI"

gene complement(join(32421..33507,36159..36931))
/gene="OJ1212_B09.8"
complement(join(32421..33507,36159..36931))
/gene="OJ1212_B09.8"
/codon_start=1
/product="receptor-like kinase"
/protein_id="BAB39439.1"
/db_xref="GI:13366218"
/translation="MEAAALSTVLVYALVADADHHVVRVQGRHQCOPFSCGHLQ
VSYPRRRGDPDRSRRCGRFSTYELDCSSDDGKATIRINTGKYVSSINTASTFWYDA
NLQDDANSSCPILPRSDQLPYVGLIGPSHSDWDLGLDMITWGVNCSQELITNSSS
KLYVYHPVDCLTTALSFVYMLPFGSLVPSVEDIEPSCGYLAMTPIGGWNVNL
EYTSAEVVKSMRGFAVKFPTPMLNWFGLIKCLNKSYSIHYOAGHTTYCIAKFISL
FVCSAMQVCVTRTVGRDPSYSQGGGYGVFLPGDVHVAIKMLDGNNGEDEF
ISEVATIGRIHNAVIRLVGFCSEEMRRAVYEMPRGSLNKYIFSSERSFSWDLNE
IALGIARGINYLHOGCEMOILHFDIKPDNILLDNEVSKVADFLAKLYPREKSFYSD
RALRGTVGYMAPEMVSRSGVTSIDKSDVYSEGMILLLEMAGRRNADPNANSRAY
PARYVDLIADQVDEISNVAADHLEKRLCLGLWCIOKSHDRPTMSEAIEMLEGD
VDALQVPLRPFCDGIGNGMPPPOVMDSYFHSSELTAFISEDDGIAELASS"
gene join(39136..39967,41208..42279)
/gene="OJ1212_B09.9"
join(39136..39967,41208..42279)
/gene="OJ1212_B09.9"
/note="probably inactive due to frameshift in CDS
pseudogene, receptor serine/threonine kinase"
/pseudo
complement(join(44067..45193,45266..45450,45538..46313))
/gene="OJ1212_B09.10"
complement(join(44067..45193,45266..45450,45538..46313))
/gene="OJ1212_B09.10"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB39440.1"
/db_xref="GI:13366219"
/translation="MAVAGAFRRSTALQALTAISVLAALLQVEGQMYGPEAAQTA
"
alignment_scores: quality: 193.00 length: 243
ratio: 1.451 gaps: 10
percent similarity: 54.733 percent identity: 25.926
alignment_block:
US-09-805-919-1 x AP003338/rev ..
Align seg 1/1 to reverse of: AP003338 from: 1 to: 112222
7 LeuphealacysvalleuleuleuValThrleuileserSerSerIysal 23
:::||||| :: ||||| :: ::::: :::::
2832 ATCTTGCTGTTCTTCTGCTTCTTGACGAGATGTGAGGGCGGCAAGG 2783
23 aglaspileserglncysvalproserSerSerIysalIegln...I 39
: ||| ||| ||||| ||| ::::: |
2782 G.....TGCATTCCTTCTCTGCTGACATCTCCAGACA 2748
39 leysPheProPheArgLeuArgThrAspProgluHisCysglYArgArg 55
:::| ||||| ||| ||||| ::::: |||||
2747 TTCTCATTCCTTCCGCGCGGAGGTGATCCACAAGATGCGGTGTTCCG 2698
56 GlyTyrGluLeuAspCysGlnAsnAsnGlnThrValPheAsnTyrIysSe 72
:::| ||||| ||||| ::::: ::::: :::::
2697 TCATATGAGTTAGACTGAGAGATAGCAAGGCTACCATTAGATCAACAC 2648
72 rargilePheaspValGlnGluIleasnTyrArgSerTyrSerIleArgL 89
: ::| ||| ::||| ||| ::::: :::::
2647 AGGACATACCTATGTAACACGACATCACTACTTCTGTTCTGCGG 2598
89 euleuAspProGlyLeuAsnAspGlnArgGluAsnCysThrValPhePro 105
:::| ||| ::||| ||| ::::: |||
2597 TTGTGATGACGAGCTTAAGATACAACAGCAGCTGC.....CCT 2557
106 AsnHisArgAlaSerTyrAspAlaMetThrSerGlnIlePheGlu..... 120
|||:::| |||:::| |||:::| |||:::|
2556 CTTCCTCGCTCGGATCAGCTTCTTCTGCGGGGGGTATCCAGGATC 2507

121TrpV 122
2506 ACATGGCGGTGGGACCTGGCCCTTGACCCCTGACCTGAGACTACATGGG 2457
122 alargValasnasnAspIleasnTyrrValasnCysLeuAlaProile... 137
2456 TT.....AGCTTTGTGAATTGTTTCACAGCGCAGTAAGG 2425
138 AsnSerSerGlnTyrlleProThrSerPheCysSerLysAsnSerThrG1 154
2424 AATAATAGTGTGTACGTTCCCTGTTGAT...TGCCCTGAGCACACAGCTCTTC 2378
154 yPheSerTyrrLeuValIleArgLulIleuGlnAlaSer..... 167
2377 GTTGTATTATGTGTGTGTTCCCTGGATAATGCCCTCCATCTGTTTCCAA 2328
168AspLeuAlaGlyGlyCysArg.....ValGluThr 177
2327 TCATCGGGAACCTTGAGACTTCATGTCGCTACCTTGCCATGATCCCTTG 2278
178 ValAlaTrpSerSerAlaProGlyIleSerSerAsnLysSerSerThrLe 194
2277 GGTGGTTGGGACTCGCCATTGCCACATAATGCAAGCTTTTCAGACATCGT 2228
194 userSerThrHisGlnGlyLeuAlaTyrrGlyPheGluLeu.....Sert 209
2227 CAGATCCATGAGGAATGGATTGCTGTTCAATTTCCCATATAATACATAGGT 2178
209 rPLysArgAsnLeuLeuCysArgAsnCys 218
2177 GGAGTAGGATTGGGCACATCAAGGATTGT 2149
seq_name: gb_pl:AP003219

seq_documentation_block: 126434 bp DNA linear PLN 03-JUL-2001
LOCUS AP003219
DEFINITION Oryza sativa genomic DNA, chromosome 1, BAC clone:OSJNB0032H19.
ACCESSION AP003219
VERSION AP003219.3 GI:14587252
KEYWORDS
SOURCE Oryza sativa (cultivar:Nipponbare) DNA, clone:OSJNB0032H19.
ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.
1 (bases 1 to 126434)
Sasaki,T., Matsumoto,T. and Yamamoto,K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
clone:OSJNB0032H19
Published Only in Database (2001) In press
JOURNAL 2 (bases 1 to 126434)
REFERENCE Sasaki,T., Matsumoto,T. and Yamamoto,K.
AUTHORS Direct Submission
TITLE Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
JOURNAL Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgrp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Jul 3, 2001 this sequence version replaced gi:13620984.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologes of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DBJ accession no. and RGP clone ID.

COMMENT
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering

almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from M13rev to -21M13 of the BAC
clone. This sequence of OSJNB0032H19 clone has an overlap with
P0439B06 clone (DBJ : AP002882) at the position 1 to 45,130 of 5'
end. The sequence of this clone starts at the position 90,862 of
P0439B06. Detailed information on overlap and assembly quality
together with annotation of this entry is available at
<http://rgrp.dna.affrc.go.jp/Genomeseq.html>.

FEATURES
source

1..126434
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="OSJNB0032H19"
complement(3679..4206)
/gene="OSJNB0032H19.1"
complement(3679..4206)
/gene="OSJNB0032H19.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB61171.1"
/db_xref="GI:14587253"
/translation="MARAILHLHSDETVPLVVIYAAYAAVALVIVSLCTSSKHGL
WKQOQRRSGSLPQLKAESAGGGSARKOLLATLSGIGKAAAMAKVSNRRSSR
QGCSDDDDDEAGGGGLAVEEEFALWRTIMGDKCRPLQPSGHIAVYDSDGNLPATTI
SKEANADAVNNTYV"

gene
CDS
complement(Join(4890..5008,6401..6539))
/gene="OSJNB0032H19.2"
complement(Join(4890..5008,6401..6539))
/gene="OSJNB0032H19.2"
/note="unknown protein"
/codon_start=1
/protein_id="BAB61172.1"
/db_xref="GI:14587254"
/translation="MQQVKRMDILPMLRDNDNVIHNGEFGVNLQVNGSSVKTSSSETL
IAEWKRLQVAVRTVVDLSVTSYTSSEKEDPILWLKFOR"
6879..7214
/gene="OSJNB0032H19.3"
6879..7214
/gene="OSJNB0032H19.3"
/note="contains EST AU082204(C52281)
unknown protein"
/codon_start=1
/protein_id="BAB61173.1"
/db_xref="GI:14587255"
/translation="MPSLRPILLAATAAAVAVSLMPAARVYADIDPSSGMPSPATTS
PEPCPTTPAGISTQPFPMWAEPSPTTTTMEPPQDPGLASACHARTAAALAWIPVFV
ESAFVLFLQ"
Join(8942..8980,9207..9524)
/gene="OSJNB0032H19.4"
Join(8942..8980,9207..9524)
/gene="OSJNB0032H19.4"
/note="contains ESTS AU101371(E20219),C99479(E20219)
unknown protein"
/codon_start=1
/protein_id="BAB61174.1"
/db_xref="GI:14587256"
/translation="MARWLQPLPCSLLRPTARRRTEAETTRRRATTRPATRRR
RGGSSRRTRCPTSRRRGRSTRTRTPASGPIRRRAAPRRGARWCSSRRRRWPEPMPCDRRR
WVINYTAFLIPFNGS"
complement(10569..10736)
/gene="OSJNB0032H19.5"
complement(10569..10736)
/gene="OSJNB0032H19.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB61175.1"
/db_xref="GI:14587257"
/translation="MELKHGISVDSNAATTDRSIECDISVILDYLLDMDAYVENAMLI
DDIFNTLLHSS"

gene
CDS

[illegible]

```

seq_documentation_block:
LOCUS      AP003219                126434 bp    DNA             linear    PLN 03-JUL-2001
DEFINITION Oryza sativa genomic DNA, chromosome 1, BAC clone:OSJNB0032H19.
ACCESSION  AP003219
VERSION    AP003219.3   GI:14587252
KEYWORDS
SOURCE
ORGANISM   Oryza sativa (cultivar:Nipponbare) DNA, clone:OSJNB0032H19.
            Oryza sativa
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
            Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE  1 (bases 1 to 126434)
AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE      Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, BAC
            clone:OSJNB0032H19
JOURNAL    Published Only in Database (2001) In press
REFERENCE  2 (bases 1 to 126434)
AUTHORS   Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE      Direct Submission
JOURNAL    Submitted (19-FEB-2001) Takuji Sasaki, National Institute of

```

COMMENT

On Jul 3, 2001 this sequence version replaced gi:13620984. Genes were predicted from the integrated results of the following: GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor (October 1998 version). The genomic sequence was searched against NCBI NonRedundant Protein database, nr (<ftp://ncbi.nlm.nih.gov/blast/db>) and the cDNA sequence database at RGP. Protein homologues of the coding regions were searched against NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent the identified cDNA sequences using BLASTN 2.0 with the corresponding DDBJ accession no. and RGP clone ID. A gene with identity or significant homology to a protein is classified based on the protein name to indicate the homology level such as same name, 'putative-' and '-like protein'. A gene without significant homology to any protein but with EST homology (covering almost the entire length of partial sequence) is classified as an 'unknown' protein. A gene predicted with a gene prediction program is classified as a 'hypothetical' protein. The orientation of the sequence is from M13rev to -21M13 of the BAC clone. This sequence of OSRNBb0032H19 clone has an overlap with P0439B06 clone (DDBJ : AF002882) at the position 1 to 45,130 of 5'

FEATURES
source

end. The sequence of this clone starts at the position 90,862 of P0439B06. Detailed information on overlap and assembly quality together with annotation of this entry is available at <http://rgp.dna.afrc.go.jp/GenomeSeq.html>.

gene /organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="OSJNB0032H19"
complement(3679..4206)
/gene="OSJNB0032H19.1"
complement(3679..4206)
/gene="OSJNB0032H19.1"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB61171.1"
/db_xref="GI:14587253"
/translation="MARAILIHLHSDETVPLLVIVYAAYAVALVIISLCTSSKHGKL
WKQOQRSGSLPOLKAESGGGGASRKOLLATLSIGKAAAMAKVSMNRSSR
QGSSDDDEAGGGIAVEEBEALMRKTIIMGDKCRPLQFSGHIAYDSGNQLPATII
SKEANADAVNNIVY"
complement(join(4890..5008,6401..6539))
/gene="OSJNB0032H19.2"
complement(join(4890..5008,6401..6539))
/gene="OSJNB0032H19.2"
/note="unknown protein"
/codon_start=1
/protein_id="BAB61172.1"
/db_xref="GI:14587254"
/translation="MQVKRRDILPMLRDNQVINGEFGQVNLQVNGSSVKTSETL
IAEWKRQVAVRTVVDLSVTSYSEKEDPLIWLKFOR"
6879..7214
/gene="OSJNB0032H19.3"
6879..7214
/gene="OSJNB0032H19.3"
/note="contains EST AU082204(C52281)
unknown protein"
/codon_start=1
/protein_id="BAB61173.1"
/db_xref="GI:14587255"
/translation="MPSLRPILIAAAVAVVSIMPARVADIDPSSGMPSPATTS
PFPECTTPAGISTQPFWABESPSTTTTTFPODPGLASACHARTALAWIPVFV
FSAFLVFLQ"
join(8942..8980,9207..9524)
/gene="OSJNB0032H19.4"
join(8942..8980,9207..9524)
/gene="OSJNB0032H19.4"
/note="contains ESTs AU101371(E20219),C99479(E20219)
unknown protein"
/codon_start=1
/protein_id="BAB61174.1"
/db_xref="GI:14587256"
/translation="MARWLQSCSILLRPTARRRTEAEATTTRRATTTRRRRR
RGGSRRTCRPTSRRRGRSTRTPASGPTRRRAAPRGARWCSSRRRWPEPWCDDRRR
MVINYAPLIPIFIMCS"
complement(10569..10736)
/gene="OSJNB0032H19.5"
complement(10569..10736)
/gene="OSJNB0032H19.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB61175.1"
/db_xref="GI:14587257"
/translation="MELKHGSIVDSNAATTDRSIECDSIVLLDVLDMDAYVEMALLI
DIFNILLHSS"
14940..15419
/gene="OSJNB0032H19.6"
14940..15419
/gene="OSJNB0032H19.6"
/note="contains EST C52872)
unknown protein"

gene
CDS

/codon_start=1
/protein_id="BAB61176.1"
/db_xref="GI:14587258"
/translation="MASYRLVLILVFSALVALAAGDTYPADCPYPCLLPPTPVTTDC
PPPPSTPSSGYSYPPPPSSSSNTPPSSSYWNTPPPQGGGGYIPYQPPAGGGGGG
GFNPAPPPPNPIVPWPYPMYYRSPSSPATAVTARGSLLASVAVVTAATAALLITVF"
complement(20579..21709)
/gene="OSJNB0032H19.7"
complement(20579..21709)
/gene="OSJNB0032H19.7"
complement(20579..21709)
/note="hypothetical protein
similar to Arabidopsis thaliana chromosome 1, F6a14.6"
/codon_start=1
/protein_id="BAB61177.1"
/db_xref="GI:14587259"
/translation="MTLAAITNPSAAAAAAYANPIVLTPGAAPPPTSALPTOI
PPSDSLSPADPALATASFTLSSTSVSPRSLSASFITLNSISLPASPNN
LPQAIRSVSPYFPALASPYASRAASLAIEDLVALLAEGLLRHPPLSLSLEADR
PELVAVYRQAADLRSSSELLATRCFISPASDAAYDAMSVKNRMKEAYAVNRCKE
KGAKKYDAMARQAALLMGYIDGFTSPVCLHYLFASENDVSVFGAAYSELDGGEV
VRMKYLYTKWIKGYGRFPEAQACPEAVGMGLELDDIVPSFRVAGALGLVFDQHFSH
LALAELKEDLKAAEMMYQDATEAESAGPILDLRMQDY"
join(21904..22051,22176..22513)
/gene="OSJNB0032H19.8"
join(21904..22051,22176..22513)
/gene="OSJNB0032H19.8"
/note="contains EST AU062688(C30153)
unknown protein"
/codon_start=1
/protein_id="BAB61178.1"
/db_xref="GI:14587260"
/translation="MVNSGETEGEPSCIGHRYTKLSPTEKPNI RHILTNOQEFDEGD
SGFI RRRPTSGEEAGSRIRVPRPHYISEKKNKWRSHGKRKPKRRADVCSVTYYART
GAEMQIESRPSSNRCAGMSYASNRIEPAALAAVASPPQCGRHRIDPTNDHLSPVSC
R"

gene
CDS

complement(join(23417..23463,23753..23936,24034..24123,
24209..24265,24570..24602,24716..24790,24885..24974,
25087..25188,25387..25457,25987..26140,26241..26317,
26821..26857))
/gene="OSJNB0032H19.9"
complement(join(23417..23463,23753..23936,24034..24123,
24209..24265,24570..24602,24716..24790,24885..24974,
25087..25188,25387..25457,25987..26140,26241..26317,
26821..26857))
/gene="OSJNB0032H19.9"
/note="contains EST AU063864(EI261)
unknown protein"
/codon_start=1
/protein_id="BAB61179.1"
/db_xref="GI:14587261"
/translation="MGLGNVAILVGSIGILVSGDAKLPSAGEVLGSAKFVKKHG
NEGKTSSNTDAHQAQLSQYNHLROEIQSLSRPVTVTNARSQPGTFITTVAAYA
GAVGYAIKWGWKLSIDLMEFYRKGLSDACNVGSQLDKVSDDVT SARKHLAGRIDRV
DISLETQEIIEGTRDEVTVHGDL SAFPQEDLOS NVLVRLESKLEVSLEYTQVPAAR
VPPAISSELVVRVSSLPOSTALPVLPTTPAEPBRAEAPOEQWGVEYRTSSCRS
EGSGRFQQRSVVTRTSSIREGSPSSNGASSSTGASTGTINTSTRFAMNASRATPW
SLMS"

gene
CDS

complement(join(28476..29162,29268..29441))
/gene="OSJNB0032H19.10"
complement(join(28476..29162,29268..29441))
/gene="OSJNB0032H19.10"
/codon_start=1
/product="putative pollen-specific membrane integral
protein"
/protein_id="BAB61180.1"
/db_xref="GI:14587262"
/translation="MTTDHACKVDVVVGNDGEHVGEQARHDLHEAAAAAADH
HAIRGLAIGFLIREVMVEGLSFLVFWSCVAALMQMYGLTFPMVCIVAMTVAEV
LSWLGPAPHNPVATTFFAAYRREPVPWKPLVYAQDIAGSLACL SVNAVBRPHDF
YGTAPVYVGHGTRLPLFMELFASVLMIVIAVATDGTAGTVGGIAIGAAGGLIVTI
GPSVGGSMNPARTLGPAIVLGRIDGVWIYVAVPAGMLYGALCNRAVRLSHRIYAFLC
GSTVGIAGSP"

complement(join(31272..32073,34221..34489))

[illegible]

LOCUS AP003250 146670 bp DNA linear PLN 01-SEP-2001
DEFINITION Oryza sativa genomic DNA, chromosome 1, PAC clone: P0443D08.
ACCESSION AP003250
VERSION AP003250.3 GI:15408719
KEYWORDS
SOURCE
ORGANISM Oryza sativa (cultivar: Nipponbare) DNA, clone: P0443D08.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzaceae; Oryza.
REFERENCE
AUTHORS 1 (bases 1 to 146670)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone: P0443D08
JOURNAL Published Only in Database (2001) In press
REFERENCE 2 (bases 1 to 146670)
Sasaki, T., Matsumoto, T. and Yamamoto, K.
AUTHORS Direct Submission
JOURNAL Submitted (19-FEB-2001) Takuji Sasaki, National Institute of
Agrobiological Sciences, Rice Genome Research Program, Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail: tsasaki@nias.affrc.go.jp, URL: http://rgp.dna.affrc.go.jp/,
Tel: 81-298-38-7441, Fax: 81-298-38-7468)
On Aug 31, 2001 this sequence version replaced gi:14522861.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from T7 to SP6 of the PAC clone.
This sequence of P0443D08 clone has an overlap with OSJNBa0083M16
(DDBJ: AP003214) at the position 1 to 624 and with P0480E02 (DDBJ:
AP002913) clone at the position 141,467 to 146,670 of 3' end. The
sequence of this clone starts at the position 138,088 of
OSJNBa0083M16 and ends at the position 5205 of P0480E02. Detailed
information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES
Source
Location/Qualifiers
1..146670
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="P0443D08"
2727..3227
/gene="P0443D08.1"
2727..3227
/gene="P0443D08.1"
/note="contains ESTs D48568(S14863), AU089765(S14863)
similar to 16.9 kda heat shock protein
unknown protein"
/codon_start=1
/protein_id="BAB64123.1"
/db_xref="GI:15408720"
/translation="MSLVKLFDTLAFDAWNPFSLFGTTVAADAWLASDTSAFANTYIE
SRETAAYVFRADLPAGYKKEEVRVEDEGNLVITGERSVRREREGQSRSHIERSCA
TEFGRRHLDDAVVDLVRASMDGMLTIVPKVYTDKQPAIAAAPVPAVPAVEAK
AIEASP"
gene complement(join(4071..4126,4283..4398,5525..5601))
/gene="P0443D08.2"
CDS complement(join(4071..4126,4283..4398,5525..5601))

/gene="P0443D08.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB64124.1"
/db_xref="GI:15408721"
/translation="MGYAWFIATINPKFDNAKFFRRVCAPOSTAGKACKLVRRHGEY
AGIQSVNPRSPDLWLSLFPSEGGVGLSATVRRYLTH"
CDS complement(6051..6500)
/gene="P0443D08.3"
complement(6051..6500)
/gene="P0443D08.3"
/note="contains EST AU093701(C63333)"
/codon_start=1
/product="putative 16.9 kda heat shock protein"
/protein_id="BAB64125.1"
/db_xref="GI:15408722"
/translation="MSLVRSNVDFEADFMDFDGVRSVLPATSDRPTAAAFANARV
DKETPESHVFKADLPGVKKEEYVEEENLVISGQRSKEKEDKNDKWHRRVERSSG
QEMRRFRLPENAKVDQVKASMEGNLVITVPKAEVKKEVKAIEISG"
complement(8660..10035)
/gene="P0443D08.4"
/note="pseudogene
similar to non-LTR retroelement reverse transcriptase"
/pseudo
complement(8660..10035)
/gene="P0443D08.4"
complement(10583..11035)
/gene="P0443D08.5"
complement(10583..11035)
/gene="P0443D08.5"
/note="contains EST AU092880(C60687)"
/codon_start=1
/product="putative 16.9 kda heat shock protein"
/protein_id="BAB64126.1"
/db_xref="GI:15408723"
/translation="MSLVRSNVDFEADFMDFDGVRSVLPATSDNDTAAAFANAR
IDWKETPESHVFKADLPGVKKEEYVEEENLVISGQRSKEKEDKNDKWHRRVERSS
GQFMRFRLPENAKVDQVKAGLENGLVITVPKAEVKKEVKAIEISG"
13625..14077
/gene="P0443D08.6"
13625..14077
/gene="P0443D08.6"
/note="contains EST AU092880(C60687)"
/codon_start=1
/product="putative 16.9 kda heat shock protein"
/protein_id="BAB64127.1"
/db_xref="GI:15408724"
/translation="MSLVRSNVDFEADFMDFDGVRSVLPATSDNDTAAAFANAR
IDWKETPESHVFKADLPGVKKEEYVEEENLVISGQRSKEKEDKNDKWHRRVERSS
GQFMRFRLPENAKVDQVKAGLENGLVITVPKAEVKKEVKAIEISG"
13625..14077
/gene="P0443D08.6"
13625..14077
/gene="P0443D08.6"
/note="contains EST AU092880(C60687)"
/codon_start=1
/product="putative 16.9 kda heat shock protein"
/protein_id="BAB64128.1"
/db_xref="GI:15408725"
/translation="MEGGHGGGWYASHFVDVEREWKRRVKSFWVLVRRLLCCRRK
SYRPDYAAGEEDDGEKSSLLTSSRSLEELLVSDADDDGAIIDAAVTCRNASLCAR
KDGAAPVAVLPBPGLHHPVMAHPAGWVTTSSGGRDGAIVQCRRRFMFGLLRRLLM
RRPWPVLVAIPE"
complement(join(19682..20619,20791..21045,21998..22714,
23306..23926,24506..25205))
/gene="P0443D08.8"
23306..23926,24506..25205)
/gene="P0443D08.8"
/note="contains ESTs AU030772(E60216), AU030771(E60216)
similar to wall-associated kinase
unknown protein"
/codon_start=1
/protein_id="BAB64129.1"

31 roserSerCysGlyAspIleGlnIleLysPheProPheArgLeuArg... 46
 36855 CGGCGAGGTGCGGGAACCTTACCATTGGCTATCCATTCTGGCTGCCGGC 36806
 47 ThrAspProGluHisCysGlyArgArgGlyTyrGluLeuAspCysGlnAs 63
 36805 AAGCATCCGCCGGAGTGTGGCTACCGAACTTTTCAGGTACAGTGGACCA 36756
 63 nasGlnThrValPheAsnTyrLysSerArgIlePheAspValGlnGlu 80
 36755 TAGAATGCTTCCCTCAAGAACGGCATCTGGACGTACCAATTCAGCGCA 36706
 80 leasnTyrArgSerTyrSerIleArgLeuLeuAspProGlyLeuAsnAsp 96
 36705 TCTTCTACCAACAACAGCTCTTCATGTGTACCAACGAGCAGCTGACGAC 36656
 97 GlnArgGluAsnCysThrValPheProAsnHisArgAlaSerTyrAspAl 113
 36655 GGCCAG....TGCGTCATCGAGAGTTTGTCAACGCCCTCTCCGACCT 36612
 113 aMetHisSerGlnIlePheGluTyrValArgValAsnAsnAspIleAsnT 130
 36611 CGGCTCAGCGCAG..TTCAAGATTAGCCCCATAAACGGAGCTCGTCT 36565
 130 YrVal...AsnCysLeuAlaProIleAsnSerSerGlnTyrIleProThr 145
 36564 TCCTCTACCAACTGC.....AGCCAGAGTGCAGCGCAGCTTCCAATC 36524
 146 SerPhe.....CysSerLysAsnSerThrGlyPheSerTyrLe 158
 36523 TCTTGGCGCCCGTGAGTGTGGCGAAATGAATCGTCCAACTCGTATGC 36474
 158 uValIleArgGluIleLeuGlnAlaSerAsp.....LeuAlaG 171
 36473 CTGGCTTCGGGGGAAGTACAGACCTGATGACGATTTCCAGACAACTGCCG 36424
 171 LysGlyCysArgValGluThrValAlaTyrSerSerAlaProGlyIleSer 187
 36423 GGAACCTGCACGGGTGTCAATGATCCCGGTGTGTGGGTACGACGGTGGGTG 36374
 188 SerAsnLysSerSerThrIleuSerSerThrHisGlnGlyLeuAlaTyrGl 204
 36373 GCGAAGAACTACGAGCGGCTC.....ATCAAGGGTGG 36342
 204 YPheGluLeuSerTyrIleLysArgAsnLeu.....LeuCysArgAsnC 218
 36341 ATTTCTCTCTCGACTACAGCGCGCGCGGGCGCGGACGATTGCGAAGATT 36292
 218 YsAspArgSerArgGlyGlyGluCysThrIleGluGluAsnSerAspArg 234
 36291 GCTCGAAGAGC...GGCGCGTGGTGTGGGTGAATGTCACTACGATGGG 36245
 235 AlaThrCysArgTyrTyrIleCysGluAspIleHisValSerLysLeuTh 251

251 rpheargcyslysvalglutyrtyr.....servaltyrvall 264
 | |||:: || | ||| :::
 36200 GTGTGTAAGTTCGTTCCGCTACGATCCCATTTCTTCCCGGTCCGTT 36151

264 euphephe 266
:::lllllll

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

Accession	Source	Size (bp)	Topology	Map Date
AP002747	Oryza sativa genomic DNA, chromosome 1	147857	linear	JAN-2001
AP002747	AP002747		PAC clone	P0698G03
AP002747.2	GI:11071975			

ORGANISM Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzaceae; Oryza.
REFERENCE 1 (bases 1 to 147857)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Oryza sativa nipponbare(GA3) genomic DNA, chromosome 1, PAC
clone:P0698G03
JOURNAL Published Only in DataBase (2000) In press
REFERENCE 2 (bases 1 to 147857)
AUTHORS Sasaki,T., Matsumoto,T. and Yamamoto,K.
TITLE Direct Submission
JOURNAL Submitted (07-JUL-2000) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/,
Tel:81-298-38-7441, Fax:81-298-38-7468)
On Nov 1, 2000 this sequence version replaced gi:9049480.
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologies of the coding regions were searched against
NCBI NonRedundant Protein database with BLASTP2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from T7 to SP6 of the PAC clone.
This sequence of P0698G03 clone has an overlap with P0463F06 (DDBJ:
AP002867) clone at the position 1 to 57,093 of 5' end and with
P0494A10 (DDBJ: AP002541) at the position 97,778 to 147,857 of 3'
end. The sequence of this clone starts at the position 87,230 of
P0463F06 and ends at the position 50,080 of P0494A10. Detailed
information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/GenomeSeq.html.
FEATURES
Source location/Qualifiers
1. 147857
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="P0698G03"
join(411. 1189,1994. 2088,2436. 3508)
/gene="P0698G03.1"
join(411. 1189,1994. 2088,2436. 3508)
/gene="P0698G03.1"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB17321.1"
/db_xref="GI:11071976"
/translation="MDFTNLLITVLLIAPLKYKSDVATASDDEDFKTCSSHRCSKH
GPEIRYPERLSTQPPSCGAPGMOISGSDTILDHPVIGSCKVTMIYRHVIMNAIPL
VDSSPHCLLHKFTISVNOSTAVYRPTLKAASLVGCSRSDIDTQNSIYVPTISLSLA
NNASQFWYLAYPEYMSILPLGCTIYSKDIPMYSYDKNGPNFDISIFETAKRIVIST
GETVFTWYSNVTISICOQEHGRCGFSSORDAFCOHSMVEIRTPSYTTLIRKL
CTCIPFAMLCIKIELVTSVSGTFIVLSLIVATALYISLSRYNEEIHLEKVEMLKTYG
TSKPTRYTFSEVKKIARFKDKLIGAGFTVYKGEILNGVPAVKMLENSVGECEFI
NEVATIGRIHANIVRLVLFQCSSEGTROALIYEFMPNESLEKTIYPHGSNISRELIYPD
KMLDIALGIARGMEYIHOQGNRIHFIDIKPHNILLDYSENPKISDFGLAKLCARDOS
IVTLTAARGTGYIAPELISRNFGAISYKSDVYFGLVLEMYSGRRNTEPTVENONE
FYFPEWIYERVMNGODLVLTMETTQGEKEMVRQALAIVALMCIQWNPKNRPSMTKVVNM
LTGRLONLQVPRKPTSSQNLVI"
gene complement(join(4690. 4841,5041. 5231,5395. 5432))
/gene="P0698G03.2"
CDS complement(join(4690. 4841,5041. 5231,5395. 5432))

/gene="P0698G03.2"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17322.1"
/db_xref="GI:11071977"
/translation="MCTPKLTQTVHRRRLTEEPRAVAAYVIGSLPSSPWARLRGAVY
NRLSGSGARRRSGEHGCGRRNEGEGRGNGAAEAASSATRGITAYVAAEHGRSCVRROP
EGEERGEKKREKEKKREKKMS"
gene complement(join(6926. 8061,8244. 8395,8496. 9262))
/gene="P0698G03.3"
complement(join(6926. 8061,8244. 8395,8496. 9262))
/gene="P0698G03.3"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB17323.1"
/db_xref="GI:11071978"
/translation="MVIPEFFRSTALOVSVLCLRFVVPDDIEMPLIHGIPACP
PTICGLSLKVASPFHRGIECGVASIELTCTDDKATIQIDNGTIYVTGIDYSDSTFWY
VDANISDSRNCPLEPRKRIPIYDDVSDDESSPYNIQVELDPASRMWSFEVNCSSK
EINNNVMYSPVACMGTSSTFVYVLTGRLSCTIENLEPSCGYIANTPLDGMGRVAPAYD
SLSYEDVVKEMRKCFQVQEPFRHHRNRSFKDLMESIPKSTGIGKISDWESLSYENF
MDCATAKIPSPYNILGIILYALMFWMTLAPPCRYLLAPLVLIIFLQKYLKRMITID
AYEKLRLMQOGLGPTRAYTDIIAITGHFRKKLGGGGYGSVYGVLLPGDLHVAIKIL
NGYSNCGNEEFISEVATIGRIHVNVRVLVFGCSSEMRALVYEMPRGSLDKHIFSS
ERRRSDKLINEIALGIARGINYLHQGDMQILHEDIKPHNILLDNFVKVADFGIAK
LIPRDKSFVSDRALRGTVGYMAPEMVSRSFGIISKSDVYFGLVLEMYVGGRRNADP
NADSSASKAYYPSWYDKLIADQVDEISNRANMHELERKLCIVGLMCIQKMSHDPRT
MSEAIEMLEGGVDALQVPLRPFCDGDGMPQPQVMDSYFHSSDLSAISEDDGIAELA
SS"
join(9938. 10001,10770. 11578)
/gene="P0698G03.4"
join(9938. 10001,10770. 11578)
/gene="P0698G03.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17324.1"
/db_xref="GI:11071979"
/translation="MICLNVYIISLSYFLNYSKFTPPPRIPRPAPFARLNEGII
PRDSLSRPFRLSLSLGCFVRIAPAILADVINGARELGRFPFESPAFSPFRRYKIYV
RASLFRFRHSPILAVGEVPAPCSSLSAVAEAREFAVAVPVLPTRRHLRLRVIA
DPVRRSASPADRRSTVIVIPNRAAFLRSGRSRRRCPCGVSRPPFAFPLLSALSG
AAPRRRSPAVTIGARARPPVGRAGEAFAALAAKWARPSAARARSTVDRAADAW
APTVDVPRAPSLG"
join(14575. 14697,15133. 15179,17078. 17656,17671. 18037,
18110. 18307)
/gene="P0698G03.5"
join(14575. 14697,15133. 15179,17078. 17656,17671. 18037,
18110. 18307)
/gene="P0698G03.5"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17325.1"
/db_xref="GI:11071980"
/translation="MESPTWPGTFGDVGPEGEYSALLLISPRGGRVAPAPAPRAERTA
TFPRACPQKWSRSGRADRVHKGWPRVSPVHGGPGAGWRSAFVWPVRSRTRG
RGRMTAGPTWQPRGARARRPRRHGRERRARAPMYTAGDHRGGAAFERAERGESER
RSSAHPGTITTTTRTAGVVESSGAARDDDDDGAPTVGERNGGADEVDDAAMPREVAP
SREEDRSDDGGEPPERKATARFGMPRATVLRHAADGVGNTAKRMEVSPSREGDRSD
VGGEKVDGDGGERKAREHESDGESEERRAATEERSTGSKYIATGGERGRGERKPR
AEISGAIEDAGELRGVNRDESEGEREGKGTEGEGITGGTGWRRGRKAMTGGSHPS
ARVAVGACQORARAGEPMGRGERREREREGLGAGPREEGGEDF"
join(21957. 22009,24616. 25161,25209. 25575,25648. 25845)
/gene="P0698G03.6"
join(21957. 22009,24616. 25161,25209. 25575,25648. 25845)
/gene="P0698G03.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17326.1"
/db_xref="GI:11071981"
/translation="MMHGMKGLLGCWRCADLCRGSSGRADRVHKGKBPVSPVHGGPG
ARSWRRSAPWVPVRSRTRGRLTAGPTWQPRGARARRPRRHGRERRARAPMTAGD
HRRGGAAPERARRGESERRSSAHPGTTTTTTRTAGVVESSGAARDDDDDGAPTGER
NGGADEVDDAAMPREVAPSRREDRSDDGGEPRRKATARGAADGVGNTAKRMEVSP

gene
CDS

SREGRSDVGPVKVDGGERGARREHESDESERAAETTERSTGSKIATGGGER
GRGERKPRAEISGAIEDAGELRGVNPDESEGEREGKGGETGEGITGCTGWRRRGCRRA
MTSGSHPSARVAGGPACQORARAGELMGRREREREGELGRAGPREEGGEDF"
join(26540.28903,28967.32695)
/gene="P0698G03.7"
join(26540.28903,28967.32695)
/gene="P0698G03.7"
/codon_start=1
/product="putative polypeptide"
/protein_id="BAB17327.1"
/db_xref="GI:11071982"

gene

/translation="MAEQAKVHDLSPSVSGDDGEPNPRRARATPPPPHONPKQGEAL
ERYEKSASAPVAGAGERRDGERRLVYGDGSTPOGALQAVGALLRHPVPVPDPESPA
QRMDDVANLVMTAOLGAGGRSATTKTSGAATTGSASSRRARAAAVARHSAATP
SSAPPTREDRHGEPDARDIERRRNSRTPRATEGASSSRVSPROGREDQPSMPACG
VGCRAVASLRNVRPFRPTTEKYDGSVNPTEFLQVYTTGIEAAGGDDRMANF
PMALKGQARGWLMLNPASVHSMEDLCQOTMNFQSTYPRGEEADLHAVQRGDESL
RSYIQRCQVRNTIPCIPAHAVIYAFRGVRRNMLEKIASKEQTTAELEFQADRVA
RKEAWTWNPSGSGVASAAPGSAQTGRDRRRKRSVHSGDEHVLAVEGPRATR
KGRPASDKKEAGTSPSERPASKWCSVHNTSLDLADCAVXKNLAERTRKWEEDRQE
RREGKSAAPVSGKRRSEAKQAPAYDIDGDDDLGFQEPGATLATVDGACAHVSHRS
FKAMRRELLAAPTHAETRRARMSVALTFTDQDHPCVARGQIAMVVSPTICNVKL
GRVLIDGGAALNLSPAAFDAIKAPGMVLOPSQPIIGVTPGHTEWGHIDLPVTEGGS
ANFETERVNFVADLSLPTNAVLYKMAVHAYALQMKMPGPGPISIRGDLK
VALACMEQRADHLAATKPEGDERLGTSAPTAPRORIATCEVPEVKEDALVSFLRA
NTDFAWRPADMPGVPEVIEHRLAVRPGARVROKVRQAPESQAFIREVARILTEA
GPIREVIHPEWLANPVVVKANGKLRMCIDYTDLNKACPKDPYPLPRIDQIVDSTAGC
DLCLFLDAYSGHQRMARDEDEKTAFTPVGYCYTSMPEGLKNAGPTFORTTRISL
GSQIGRNVEAYVDDLVTYTRNQETLLSLDAETENLRSARIKLNDCQVFGVPAGKLL
GFLVSARGIEANPEKIRAIERMSPSKLRDVOCTGCMALSRFTSRLGEKALPLFKL
LKRSQPTWTEEAHALTQKAYLSSPVLVAEPNEPLLLYLAATPVVSAALVER
DENNPHEVPHSVLTMPGSESGEAPESDGLRLTTGVPPLPACPTMPCOF
KATAGRPHLTPEDPEANPVLTRPGKEQGEAPENGGORPLTTGVPPLPACPTMPCOF
DPDGPPEATTEGRPLSSGDPVEVIGTEDKCAKRGHLINEBPGDTPASKEDRPHRVQR
VYEVSEALRDAKTRYPQAQKMLYALMASRKLHAYFOAHRVTVVSYPPLGQILYNREG
TGRVYKWAIELSEFDLHEPRHAIKSOTLADFEVAEMTPAPETVSIPEASTNSQLPHT
THWYQFDSLSLQAGAVTLTSPGDLRLYLRLDFRATNNMAIEGGLLAGLRVAA
GLGIRRLVLDGDSQLVVQVCKEYRCSDPQMDAYVHQRMERHEDGIELRHVPRDN
MIADLSRLASSRAQTPGAFEEERLTQPSARPDLETDAIDPRPPVGLASGPEGES
APSSRLIAMISEIQYLTDLKLPEDREGSERVQIRSKRYVLVEGLYRRAANGILK
CIPOEGVELLADIEHEGCAHASRDLVGAFRQGFYPTALNMAIEGGLLAGLRVAA
HAKOIHQPAQALQIITPLSMFPAWGLDILGPRKAPGGFEYLVAIDKFTKPEAYPV
VKIDKHSALKEIKGITAREFVNRIITDNGQFTSELFGDYCEDMKLGFASPAHPR
SNGOVERANAEILKGLKTFNLIKKGDSNTEELPAVLMANRTFSRATGETPFLV
YGAEAVLPSELTLRSPRATMYCEADQDLRRDLDLLEERRRRAARAARYOOSLRRY
HQRHVRARSLCVDLVLARVOTRAGLSKLSPMWGPYRVIGVPRPSVRLATGDSTEL
PNPWNIEHLRRFYP"
complement(join(33311.34366,34450.35245,35345.36603))

alignment_scores: Quality: 167.50 Length: 238
 Ratio: 1.279 Gaps: 10
Percent Similarity: 55.042 Percent Identity: 26.471

alignment_block:
US-09-805-919-1 x AP002747/rev ..

Align seg 1/1 to reverse of: AP002747 from: 1 to: 147857

2 LeuthrArgglyLeuLeuPheAlaCysValleuLeuValthrLeuI1 18
||| :::||||| ::
72773 TTGCAGACAGCTCTAGTATCTCTTCTTGTGCTGTGCAGCAGATGT 72724
18 eSerSerSerLySAlaGlnaspIleSerGlnCysValProSerSerCysG 35
::: ::: ||| ||| ||| |||
72723 TGGAGGAGGGCAAGCAACAC..CACTACTGTCTCCTCCTCCTCCTGTG 72677
35 LyaspIle..GlnIleLySpheProPheArgLeuArgThrAspProGlu 50
||||| ::: ||| ||| ||| ||| |||
72676 GTGATCTTCACACACATATTCCTTCCTTCCTCAAGGCGATTCGCCGC 72627
51 HisCys.....GlyArgArgglyTyrGlnLeuAspCysGlnAsn.. 63
||| ||| ||| ||| ||| ||| |||

72626 GATTGGCTTGCTGCTCCCGCTCCATGATGACCTGTCTGCAGCAACAC 72577

64 .AsnGlnThrValPheAsnTyrLySerArgIlePheAspValGlnGlu 80
::: ::: ||| ||| ||| ||| |||
72576 TGGCAGGCCACCATTACATCAACACAGACATACTACGTGAGTAGCA 72527

80 LeasNtyrArgSerTyrSerIleArgLeuLeuAspProGlyLeuAsnAsp 96
||| ::: ||| ||| ||| ||| |||
72526 TCAACTACACCGATCTGTCTCTTGTGTCACCGACGCCACCATGCAGAT 72477

97 GlnArgGluAsnCysThrValPheProAsnHisArgAlaSerTyr..... 111
::: ||| ||| ||| ||| ||| |||
72476 GATGATACAGAGCTGCCCTCTTCTCCATCCGATCACCCCTCCTTACAGTAC 72427

112 113
72426 TGACGATCTGTTGGACCGCTACTACAACATGATTCATTATGGCTTTCGATTT 72377

113 LamethrSerGlnIlePheGluTyrValArgValAsnAspIleAsn 129
||| ||| ||| ||| ||| ||| |||
72376 TGGCTACTTCTTCAGACAGTGTGGCA.....TGC 72345

130 TyrValAsnCysLeuAlaProIle.....AsnSerSerGlnTyrI1 143
::: ||| ||| ||| ||| ||| |||
72344 TTTGTGAATTTGTTCCAGACCAATACACAGATACTACTAGTCCATGGTAGAG 72295

143 eProthrSerPheCysSerLySAsnSerThrGlyPheSerTyrLeuValI 160
||| ||| ||| ||| ||| ||| |||
72294 GCCCGTCTCT..TGCCGTGCTTACCAACAATTCCTTTGTTATCTCAACG 72248

160 leargGluIleLeuGlnAlaSerAspLeuAlaGlyGlyCysArgValGlu 176
72247 AGTTTATGCGCTGTACAATTGAGAGAGCTTCAACCTTCTTGTGATACTTTC 72198

177 ThrValAlaTyrSerAlaProGlyIleSer.....SerAs 189.
::: ||| ||| ||| ||| ||| |||
72197 TCCATGATTCCTTGTATACCAACATATATACGCTTCACAGCTACAAAA 72148

189 nLySerSerThrLeuSerSerThrHisGlnGlyLeuAlaTyrGlyPhe 205
||| ||| ||| ||| ||| ||| |||
72147 TGCAAGTATACTGATATCATGATTCATAGGGAAGGGCTTTCGTGTTC 72098

206 GluLeuSerTrp 209
::: ||| ||| ||| ||| ||| |||
72097 GATTCCCAATGG 72086

seq_name: gb_pl:AF100765

seq_documentation_block:

LOCUS AF100765 3067 bp DNA linear PLN 02-AUG-1999
DEFINITION Oryza sativa receptor-like kinase (BAR1) gene, complete cds.
ACCESSION AF100765
VERSION AF100765.1 GI:5669662
KEYWORDS

SOURCE
ORGANISM Oryza sativa.

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Ehrhartoideae; Oryzeae; Oryza.

REFERENCE
AUTHORS Feuillet,C. and Keller,B.
TITLE 1 (bases 1 to 3067)
High gene density is conserved at syntenic loci of small and large
grass genomes

JOURNAL Proc. Natl. Acad. Sci. U.S.A. 96 (14), 8265-8270 (1999)
MEDLINE 99324224
2 (bases 1 to 3067)

REFERENCE
AUTHORS Feuillet,C. and Keller,B.
TITLE Direct Submission
JOURNAL Submitted (23-OCT-1998) Plant Biology, University of Zurich,
Zollikerstr. 107, Zurich 8008, Switzerland

FEATURES
source location/Qualifiers
1..3067
/organism="Oryza sativa"

```

/db_xref="taxon:4530"
/chromosome="1"
<331..1037
/gene="8ARK1"
join(331..1037,1153..1310,1386..2533)
/gene="8ARK1"
/note="similar to wheat ARK1AS"
/codon_start=1
/product="receptor-like kinase"
/protein_id="AAD46415.1"
/db_xref="GI:5669663"
translation="MAIHGVFALLLSAVISVLCILAVLPDAGGRHNRHCPPFTC
GNLSSVSPPRRRGDPSECGVLSEYELNCAADKATIQINDGYVTGINTDSTELLVD
ADISDSPSSCPPLPRMNSLRPREFQVLWTRTASFVNCSOEIKNSWMITMPVACLSTR
SFVVYLTONPALIQSLSPSCGYLAMTPGGRSAPVPNESLSIADVVKSMRKGRFR
VOFFRYSFKQCPMDSIPTGKTNWEGTII SADTNFQLCI PVHQIPSRYPMLGLIMLY
IYAFAQIWKMIAATCRFLTLPLAVLVFLAHKWTKRTVIDAVEKFLRMQMUGPTRYA
YTDITAITGHFRREKLGGGYSVYKGVLPLDLHAVKMLNGSNSNCEEFISEVATI
GMHHNVNVRIVGFCSSEMRALVEHMRPSLDKIYLSERSFSWDKINIEIALGIAR
GINYLHQGDMOILHFEDIKPHNILLDNEPVKVADEFGLAKLYPRDKSFPDRALGTV
GYMAPEMVSRSGVISSKSDVSYFGMLLEMGAGRNDVPNADSKSAIYPSWVYDQ
LIADQEDDELISVANMHLERLKICLVGLWCIOKMSHDRPTMSEAITEMEGGDVALQVP
PRPFCDGDGDTGNGMPPPQVMDSYFHSSKLTALAISEEGDI AKLASS"
<331..>2533
/gene="8ARK1"
join(<331..1037,1153..1310,1386..>2533)
/gene="8ARK1"
/product="receptor-like kinase"
1153..1310
/gene="8ARK1"
1386..>2533
/gene="8ARK1"

BASE COUNT      819 a        648 c        707 g       891 t         2 others
ORIGIN

alignment_scores:
    Quality:   166.00           Length:   228
    Ratio:     1.221            Gaps:    10
Percent Similarity: 59.649      Percent Identity: 23.246

alignment_block:
US-09-805-919-1 x AF100765 ..

Align seg 1/1   to: AF100765   from: 1   to: 3067

1 MetLeuThrArgglyLeuleuPhealacysValleuleuValThrle 17
:::| | :::::::::: | | | | | | | | | | :
361 CTACTGCTATCTGCGGTAAATTTCCTCCCTGTTGCTGCACATTCTTGP 410
17 uileSerSerSerlyAlaglnAspileSerGlnCysvalProserSerc 34
: ::::::::::: : | | | | | | | | | | :
411 CCCAGATGCCAGGTGGCGGCATCATCGCATGCGTGCCCTCCGTTCACT 460
34 ysgLyaspIle..glnIlelySpheProPheArgLeuArgThrasPro 49
|| || | | | | | | | | | | | | | | | | | | | | | | | | | |
461 GTGGTAATCTTAGACAGCGTGTCGTCCTTCCTCCGCCGCGAGGTACCCA 510
50 GluHisCysglyArgArgglyTYTrgluLeuaspcCysGlnasnaInclnTh 66
::::::::::| | | | | | | | | | | | | | | | | | | | | | |
511 TCTGAGTGGCGAGTCTCTGTCTGACGAGCTCAATTGTGCCGATGACAAGC 560
66 rValpheasnTyrlsSerArgllePheaspValgInguilleasnTyra 83
::::::::::| | | | | | | | | | | | | | | | | | | | | | |
561 TACAATTCAGATTGACACACGACACATACTATGTGACGGGCATCACTACA 610
83 rgSerTySerlleArgleuleuaspproglyleuasnaspglInarglu 99
::: :::::| | | | | | | | | | | | | | | | | | | | | | |
611 CTGATTCCTACCTCTGTGATCGTTGATGCTGACATCTCGGATTCACTAGC 660
100 AsnCysThValalpHeProasnHisArglaSerTyrraspAlamethrse 116
::::| | | | | | | | | | | | | | | | | | | | | | | |
661 AGTtGCCCTCTT.....CTCGATGAATTCCTCGAGAGA 695

```

```

116 rglnilepheglu.....TrpValArgValAlaAsnAsnSpIleasntYrv 131  

   : |||::: ||:::||:::  

696 GCCCAGGTTCCAGGTTCATGACTAGGATA.....GCTAGCTTGG 736  

131 alaSnCySLeuAlAProIleaSnSerS.....GIntYrILEPro 144  

   |||| |||::: ||  

737 TCAATTGTTGCAGGAATAAAGAATAGTAGCTGATTATGTACATGCC 786  

145 ThrSerPheCySerLySaSnSerThrGLyPheSerTyRLeuValILear 161  

   :::: ||| ::::||| ::||| |||::::  

787 GTCGCT...TGCGTAGCACACCAGATCCCTCGTTTA GTTAACTGG 833  

161 gGU.....lIleucInAlaserASplEuAlaglYLcysArgv 175  

   :: ::::|||::: ::: :::  

834 CCAAATCCCTGCTCTATTCAGAGCCTTGAGCCCTCTGTGGTACCTAG 883  

175 alGUThrValAlatrp.....SerSerAlapro.....gly 185  

   :: ||| ||| |||||  

884 CCATGACTCCTTGGGTGTGGGACAGTAGTACACCAATAATGAAGC 933  

186 lIeSerSerASnlYSerSerThrLLeuSerSerThrHisgIngly.... 200  

   ::||| :::: :::::||| ::::|||  

934 TTAgCTTAAGCTATGCAGACGTAGTAAATCTATGAGAGAAAGATTcag 983  

201 .LeuAlAtyrGlYPhegluleuSertrplysArg 211  

   :: ::: |||::: |||:::|||:::  

984 AGTCCAATTTCCTTCAGATATAGCTTTAAGCCAA 1017  

seq_name: gb_pl:AF238475  

seq_documentation_block:  

LOCUS AF238475 3999 bp DNA linear PLN 21-JUN-2000  

DEFINITION Oryza sativa receptor-like kinase (RLG17) gene, complete cds.  

ACCESSION AF238475  

VERSION AF238475.1 GI:8575483  

KEYWORDS  

SOURCE  

ORGANISM  

Oryza sativa.  

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  

Eurharoidae; Oryzaceae; Oryza.  

REFERENCE  

AUTHORS Yun,C.-H., Lee,G.-R. and Kim,H.-I.  

TITLE Molecular analysis of receptor-like kinase in rice  

JOURNAL Unpublished  

REFERENCE 2 (bases 1 to 3999)  

AUTHORS Yun,C.-H., Lee,G.-R. and Kim,H.-I.  

TITLE Direct Submission  

JOURNAL Submitted (24-FEB-2000) CytoGenetics, National Institute of  

Agricultural and Technology, Seodundong 249, Suwon 441-707, Korea  

FEATURES  

source  

1..3999  

/oranism="Oryza sativa"  

/variety="Nipponbare"  

/db_xref="taxon:4530"  

/chromosome="1"  

join(<787..1502,1673..1803,1915..>3020)  

/gene="RLG17"  

/product="receptor-like kinase"  

<787..>3020  

/gene="RLG17"  

join(787..1502,1673..1803,1915..3020)  

/gene="RLG17"  

/note="protein kinase"  

/codon_start=1  

/product="receptor-like kinase"  

/protein_id="AAF78019.1"  

/db_xref="GI:8575484"  

/translaton="MRVIVLVCLGVLPDAAGRHRHRCDFPTGHLSDVSEFPF  

RRGDPEEGVOSEYEELTKADDKATIQIDKTYSVSDINSGDSTLWVDASFDSRSSC  

LLPRWNPLLRDRBLRGASHHIETLAIPAVGVTWASFVNCSQEIRNSNMMPVACLSTS  

RSFYVFETGGQSAYIONLEPSGCVLATTPLGSKLNSTSAIQNVSYODVVKLMVTGFA
```



```

7556 AGTGACATCAACTATGATGATTTCTACCTGTGGGTGCTGCATGCCAGCTT 7507
      :|||:|||||:  ::  :|||:  :|||:  :|||:
106 uAsnAspGlnArgGlnAsnCysThrValPheProAsn..... 106
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7506 CTGGATTACAGTAGAGTTGC...CTTCTCTCCTCGCTGGAACCCCTTC 7460
      :|||:|||||:  ::  :|||:  :|||:  :|||:
107 ..HisArgAlaSerTyrAspAlaMetThrSerGlnIlePheGlu..... 120
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7459 TTCGTGATCCAGCGCTCCAGGCCAAGTCACACCACATCATTTGAATTGGCA 7410
      :|||:|||||:  ::  :|||:  :|||:  :|||:
121 .....TTPValArgValAsnAsnAspIleAsnTyrVa 131
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7409 CCTCCAGTAGGGGTACTCGGGCT.....AGCTTTGT 7378
      :|||:|||||:  ::  :|||:  :|||:  :|||:
131 lAsnCysLeuAlaProIleAsnSerSer.....GlnTyrIleProThrs 146
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7377 CAATTGTTCCGAGAAATAGAAGACAGCAGTTGGATGTACATGCCCGTCG 7328
      :|||:|||||:  ::  :|||:  :|||:  :|||:
146 erPheCysSerLysAsnSerThrGlyPheSerTyrLeuValIle..Arg 161
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7327 CT..TGCTGAGCACCAGAGATCCTTCGTTACGTGTTTACTGGCCAA 7281
      :|||:|||||:  ::  :|||:  :|||:  :|||:
162 GluIleLeuGlnAlaSerAspLeuAlaGlyGlyCysArgValGluThrVa 178
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7280 CAATCTGCTTATATTCAGAACCTTGAGCCCTTCCTGTGGGTAC.....CT 7237
      :|||:|||||:  ::  :|||:  :|||:  :|||:
178 lAlaTrpSerSerAlaProGlyIleSerSerAsnLysSerThrLeus 195
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7236 GGCCACGACTCCTTTGGTGGCAGCAAGTTGAACAGCAGCGCGCTG. 7188
      :|||:|||||:  ::  :|||:  :|||:  :|||:
195 erSerThrHisGlnGlyLeuAlaTyr 203
      :|||:|||||:  ::  :|||:  :|||:  :|||:
7187 .....CAGAAATGTAAAGCTAT 7173
      :|||:|||||:  ::  :|||:  :|||:  :|||:

seq_name: gb_pl:AC006418

seq_documentation_block:
LOCUS AC006418 67730 bp DNA linear PLN 05-APR-2000
DEFINITION Arabidopsis thaliana chromosome II section 247 of 255 of the
complete sequence. Sequence from clones F13A10.
ACCESSION AC006418 AE002093
VERSION AC006418.3 GI:6598561
KEYWORDS HTG.
SOURCE thale cress.
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
Rosidae; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 67730)
Lin,X., Kaul,S., Rounsley,S.D., Shea,T.P., Benito,M.-I., Town,C.D.,
Fuji,C.Y., Mason,T.M., Bowman,C.L., Barnstead,M.E.,
Feldblum,T.V., Buell,C.R., Ketchum,K.A., Lee,J.J., Ronning,C.M.,
Koo,H., Moffat,K.S., Cronin,L.A., Shen,M., VanAken,S.E., Umayam,L.,
Tallon,L.J., Gill,J.E., Adams,M.D., Carrera,A.J., Creasy,T.H.,
Goodman,H.M., Somerville,C.R., Copenhaver,G.P., Preuss,D.,
Nierman,W.C., White,O., Eisen,J.A., Salzberg,S.L., Fraser,C.M. and
Venter,J.C.
Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana
JOURNAL Nature 402 (6763), 761-768 (1999)
MEDLINE 20083487
PUBMED 10617197
2 (bases 1 to 67730)
Lin,X.
Direct Submission
Submitted (09-MAR-2000) The Institute for Genomic Research, 9712
Medical Center Dr., Rockville, MD 20850, USA
On Dec 17, 1999 this sequence version replaced gi:4415928.
The sequence and annotation of chromosome 2 were merged from those
of the individual clones on this chromosome after removing
overlaps. For detailed information, please see the TIGR web site
(http://www.tigr.org/tdb/at/at.html).
COMMENT
JOURNAL
TITLE
AUTHORS
REFERENCE

```

Genes were identified by a combination of three methods: Gene prediction programs including GRLT (<ftp://arthur.epm.ornl.gov/pub/xgrail>), GeneFinder (Phil Green, University of Washington), GENSCAN (Chris Burge, <http://genomic.stanford.edu/GENSCAN.html>), and NetPlantGene (<http://www.cbs.dtu.dk/services/NetGene2/>), searches of the complete sequence against a peptide database and plant EST databases at TIGR, and manual curations based on those analyses. Annotated genes are named to indicate the level of evidence for their annotation. Genes with similarity to other proteins are named after the database hits. Genes without significant peptide similarity but with EST similarity are named as 'unknown' proteins. Genes without protein or EST similarity, that are predicted by two or more gene prediction programs over most of their length are annotated as 'hypothetical' proteins. Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, <http://genome.wustl.edu/eddy/tRNAscan-SE/>). Simple repeats were identified by RepeatMasker (Arian Smit, <http://ftp.genome.washington.edu/RM/RepeatMasker.html>). Genes are numbered from the top to bottom of the chromosome.

We thank the GSH/WashU/ABI consortium for sequencing BAC clones F6P23, F5J6, T17A5, and T13L16, the ESSA group for sequencing clone F13D4, and Scott Jackson, Jiming Jiang, Klaus Meyer, Eric Richards and Satoshi Tabata for helpful assistance. In addition, we would like to thank the TIGR Bioinformatics Department, especially Lixin Zhou, Hanif Khataf, Michael E. Heaney, Lily Fu, Feng Liang, Jeremy Peterson, Michael Holmes, and Delwood Richardson for software and database support.

This work was supported by the National Science Foundation, Department of Energy and the US Department of Agriculture.

Address all correspondence to: at@tlgr.org.

```

FEATURES
    source
        Location/Qualifiers
            1..67730
                /organism="Arabidopsis thaliana"
                /cultivar="Columbia"
                /db_xref="taxon:3702"
                /chromosome="I1"
    misc_feature
        1..>67730
            /note="Sequence from clone F13A10"
    mRNA
        complement(join(<1047..1322,1409..1518,1595..1997,
        2074..2431,2503..2828,2916..>3029))
            /gene="At2g46480"
            complement(<1047..>3029)
            /gene="At2g46480"
            /note="F13A10.1; predicted by genscan"
            complement(join(1047..1322,1409..1518,1595..1997,
            2074..2431,2503..2828,2916..3029))
            /gene="At2g46480"
            /note="hypothetical protein"
            /codon_start=1
            /protein_id="AAD20159.1"
            /db_xref="GI:4415929"
            /translation="MTDACCCKGNEDKMYPRFGHTWIGKAFNDTPEMLHSLRQEK
            RLERANELMNDSDLOKETAAARSRSYSDAPLGNYYTIWKNEYRGKSFEDMLRMOD
            QIIMARVYSGLAKFTNNLALHQEIETQLMKLAWEEESTDIDQEQRYLDSIRDMQIILA
            RAHQQLYECKLVYTNKLRAMLQTVDELENEQTYITFILQLASKALPDATICLTRLNLI
            EYHLLPLPMRNFPRENLENPKLYHYALESDNYLAASVVNNTVMNAQDSPRHVPHLV
            TDKLNFGAMSMWFLNPPGEATIHVORFEDFTWLNSYSPLISOLESAAMKFEYKTA
            RSESVESGSENLYKRYPKYMSMLNHLRYIIPRFPKLEKILFYDDVYVQKDLTPLMS
            IDLKGKVENENDEPKFCGWAYGMNIFDLKKNKNITETYYHFWQNLNENRTLMKGLTLP
            PGLITFYNLTPQLQRKWHLLGLGYDKGIDVYKKIERSAVIHYNGHMKPWTENGISKYQP
            YMTKYTNEDHPYIFTCRLEF"
    mRNA
        4060..4696
            /gene="At2g46490"
        4060..4696
            /gene="At2g46490"
        /note="F13A10.2"
        4100..4504
            /gene="At2g46490"
    CDS

```


JOURNAL

COMMENT

Submitted (13-SEP-2000) Takuji Sasaki, National Institute of
Agrobiological Resources, Rice Genome Research Program; Kannondai
2-1-2, Tsukuba, Ibaraki 305-8602, Japan
(E-mail:tsasaki@abr.affrc.go.jp, URL:http://rgp.dna.affrc.go.jp/
Tel:81-298-38-7441, Fax:81-298-38-7468)
Genes were predicted from the integrated results of the following:
GENSCAN1.0, BLASTN2.0, BLASTX2.0 as well as SplicePredictor
(October 1998 version). The genomic sequence was searched against
NCBI NonRedundant Protein database, nr
(ftp://ncbi.nlm.nih.gov/blast/db) and the cDNA sequence database at
RGP. Protein homologes of the coding regions were searched against
NCBI NonRedundant protein database with BLASTp2.0. ESTs represent
the identified cDNA sequences using BLASTN 2.0 with the
corresponding DDBJ accession no. and RGP clone ID.
A gene with identity or significant homology to a protein is
classified based on the protein name to indicate the homology level
such as same name, 'putative-' and '-like protein'. A gene without
significant homology to any protein but with EST homology (covering
almost the entire length of partial sequence) is classified as an
'unknown' protein. A gene predicted with a gene prediction program
is classified as a 'hypothetical' protein.
The orientation of the sequence is from SP6 to T7 of the PAC clone.
Detailed information on overlap and assembly quality together with
annotation of this entry is available at
http://rgp.dna.affrc.go.jp/Genomeseq.html.

FEATURES

source

1. 144322
/organism="Oryza sativa"
/cultivar="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
/clone="P0463F06"
join(3116..3947,5188..6259)
/gene="P0463F06.1"
join(3116..3947,5188..6259)
/gene="P0463F06.1"
/note="probably inactive due to frameshift in CDS
pseudogene, receptor serine/threonine kinase"
/codon_start=1
/pseudo
complement(join(8047..9173,9246..9430,9518..10293))
/gene="P0463F06.2"
complement(join(8047..9173,9246..9430,9518..10293))
/gene="P0463F06.2"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB17115.1"
/db_xref="GI:11034591"

gene
CDS

/translation="MAVAGAFRRSTSTALQALTLASVLAALLQVEGQMPPEEAQTA
QCPPFSWGLKNVSPRRRRGDDPDCGKSHLLVCSDTKATIRINNATYYVTRIHS
LFYVVDLDLNSHNSCP.LPRQHPHERISMAP.TSDGVQVIELAFTNRWAAEFVNC
SKEITRNVRKPVACLS.TSHSFYVLTGSESETVYESLEPCGYLGMTPLDDRGA
NASLSYADVLKYMKGFAIGISSSGTYHEDFRACLAESVSIFQTMVPSDDIYIKNR
LIAILLAYGPFWQCIMAPVKLPRAVAYEFATYGTGPAISILTWAVLCRFLAPLVLT
FLSHKYWKARITIDAVEKFLMOEMLSPMRGYTDIIAITSFRDLRGQSGYGVYK
VLLPGNIHVAVKMLNNGNSNGEEFISEVSTIGRIHVNVRVLVGCSEEMRALVYE
YMRGSLDKYIFSSDSFSRDKLINEIALGARGINYLHOGCDLQILHEDIKPHNILD
KNVPKVADEGLARLYPRDKSFVPVSAARGTVGIAPEMISRSFGVYISSKSDVYSGM
LLEMGARRNADPNMENSQAAYPSRVYRLTROETGEITAAADMELEKLCIYGL
WCIQMRSCDRPMMEVYIEMLEGVDCIQIPPRPFCDDDYIPAMESLYLSELELAI
SEKDEESITELN"
join(11349..11391,11477..12273,12952..14085)
/gene="P0463F06.3"
join(11349..11391,11477..12273,12952..14085)
/gene="P0463F06.3"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB17116.1"
/db_xref="GI:11034592"
/translation="MTRTIPILQRLKQIVLLSLINPEYSTKAWDEEGFLKSCSSQR
CSKHGPDIREPERLSTHDPSCSGHDTILDHPVLGSGCKVTAIYYRYRVIN
AIIIVESSSNCP.LQKLASTNVSTAVYEPQKKGATIVGCSKDSIATNQDGIIVGRSCL
NLSSHGSQLWYLVPPRRYMSALPDCITVAKGIPIPYNYDRNGPNENQYLDISNLKEK

gene
CDS

ANKVINGETAFTWHLNKITNACQGCENGNHCGFSSQROAFQCOHGLILIEFFHK
GILQILAHLLILLMEQNTFIYFATASSATLVLLMVGTITYLSLTRYNEIHLK
VEMFLKTYGTSPRTVRSFSEVKITRRREKIGHGXYGNVYKGLPNQVAVKMLEN
SRDGQGEINEVATIGRIHNAVLRLLGFCSEGRITLLIYEFMPNDSLEKIFYSQPN
VSRREFLPDKMLDLSIGARGMEYLHOGCNQRIREDIKPONILDYSFNPKISDFGL
AKLCARQOSIVTLTARGTMGYIAPELYSRNFQVYSKSDVYSGFMLVLEMVSGRRNL
DPGIDNOHEVYFLEMYEKVFTGQNLIGTEMTODEKYYKVKRLAIVALMCIQWPNKRN
PSTQVAVNMLTGRLODQMPKPFVSSASESIP"
complement(join(14626..14732,14821..15151))
/gene="P0463F06.4"
complement(join(14626..14732,14821..15151))
/gene="P0463F06.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17117.1"
/db_xref="GI:11034593"
/translation="MCQLGDGSRILFWEDCWMQEGGIRAIAPSIFEHYVRIRKRTV
AEALNKAWIRDISGALGVQAIREYLKLMSLQNAACNPSEQDVLCWKWDSSGQYSSK
SAYRALFLDAGRIDYVVAVYHLLTDAYFVINTRYLLITFLLA"
16173..17087
/gene="P0463F06.5"
16173..17087
/gene="P0463F06.5"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB17118.1"
/db_xref="GI:11034594"

gene
CDS

/translation="MATALYLSLKSRYDEEVLKYMELRTYGTSKPRYSFSDVKRI
TRCFKEQLGQGGFSGVNIVRLGFCSEGRHILYEFMPNESLEKITYFHDPTSQEL
LAPKMLDIALGIARGMEYLNHOGCNQRIHFDIKPHNILLDYNFNKISDFGLAKICA
RDQSVITLTKARGTMGYIAPELYSRNFGEISYKSDVYSGFMLVLEMVSGRRNSDPIE
DQNEVYFLEMYEKVITGODEFLVSGEMTEEDRLKVRQMALVALWCIQWPNRNPMSMTK
VVMNLTRQONIQVPPKPFVSYSHAVP"
complement(join(17879..18106,18186..18372,18654..18962,
19113..19291,19329..19461,19516..19624,20367..20631))
/gene="P0463F06.6"
complement(join(17879..18106,18186..18372,18654..18962,
19113..19291,19329..19461,19516..19624,20367..20631))
/gene="P0463F06.6"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17119.1"
/db_xref="GI:11034595"
/translation="MGCGQEQRRSSCSRSSPPASSGCWSTSSMSATGVAVGEYRAGG
LVDEREQPYTELARLHCRRPSPILPPEFLVVPSTPYSYVGELAPGEDVGLLPSAGQ
GGRGRRRRRPVTGDTMAAGCAAEYVADPEIEGHELVGQMRPAPGEHAGCKRAGV
EAAEDEEQASGRRLPLGLKLIAVVDVQDTKLDNQDLSYGGKGYFIRDIEPNLHQ
NONHNQINQSMTMTLSSVRLRDVAMASNSRSTLPSAAAIYARFLTRPAAADPNAA
VNPAAVLAISAADLRAIGVSAKAAYLHDLAGRAAGELSESAAVADDEAAPASRQ
PTPLRSPAYPNEKSERERERERKGEERVMTWPMADMGPRGSHADSAVTSKTKMKTTEG
PRLRLRCFTFSIDPFPPLATNPSSPAIPPTPGHLCFPGDVSALRSAVNARCPOYASP
SPCVLVRPEPQIQAPEPD"
join(22482..22821,23165..23209,24203..24528)
/gene="P0463F06.7"
join(22482..22821,23165..23209,24203..24528)
/gene="P0463F06.7"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17120.1"
/db_xref="GI:11034596"
/translation="MPARCSHRPRLSSAGRADHMLSSMPDVTMATDPTIAYIMDE
QOEIKFKTSKCWNPIRPPARLITNSGRQIYIRPPSAREYILMESSWSPISNGFSLISK
SHLIWPQSKKOGIELICTGKINPNRVAVPCRAGPRAPAAQARHGYSGRAGTARYCS
RPCRAVPATGSAVPAARLARVPWPLWLSSHQCAAPQPSLTMTFNNAVNRNVCRTTGP
RQIGCDQSYIGLSPVTS"
join(29524..29612,29887..29965,30113..30169,30207..30533)
/gene="P0463F06.8"
join(29524..29612,29887..29965,30113..30169,30207..30533)
/gene="P0463F06.8"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17121.1"
/db_xref="GI:11034597"

gene
CDS

ANKVINGETAFTWHLNKITNACQGCENGNHCGFSSQROAFQCOHGLILIEFFHK
GILQILAHLLILLMEQNTFIYFATASSATLVLLMVGTITYLSLTRYNEIHLK
VEMFLKTYGTSPRTVRSFSEVKITRRREKIGHGXYGNVYKGLPNQVAVKMLEN
SRDGQGEINEVATIGRIHNAVLRLLGFCSEGRITLLIYEFMPNDSLEKIFYSQPN
VSRREFLPDKMLDLSIGARGMEYLHOGCNQRIREDIKPONILDYSFNPKISDFGL
AKLCARQOSIVTLTARGTMGYIAPELYSRNFQVYSKSDVYSGFMLVLEMVSGRRNL
DPGIDNOHEVYFLEMYEKVFTGQNLIGTEMTODEKYYKVKRLAIVALMCIQWPNKRN
PSTQVAVNMLTGRLODQMPKPFVSSASESIP"
complement(join(14626..14732,14821..15151))
/gene="P0463F06.4"
complement(join(14626..14732,14821..15151))
/gene="P0463F06.4"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17117.1"
/db_xref="GI:11034593"
/translation="MCQLGDGSRILFWEDCWMQEGGIRAIAPSIFEHYVRIRKRTV
AEALNKAWIRDISGALGVQAIREYLKLMSLQNAACNPSEQDVLCWKWDSSGQYSSK
SAYRALFLDAGRIDYVVAVYHLLTDAYFVINTRYLLITFLLA"
16173..17087
/gene="P0463F06.5"
16173..17087
/gene="P0463F06.5"
/codon_start=1
/product="putative receptor kinase"
/protein_id="BAB17118.1"
/db_xref="GI:11034594"


```
/translation="MKLIVHPIIMDKYPLHAVVPRERROPLSLSTLSDFHOSSIVP
EKEHIKILPRSSQLHSRPRPRPTSSGKITRLNALHVAIEIVNSDALOQPA
RADTEHASGHSYCGPLHGVPILKDNIVKRDRLNTAGSFALLGSSVVCRDAGVTAGLRA
AAAAIILSKANPVPNGWSARW"
join(34905.34955,35683.36040,36863.36942)
/gene="P0463F06.9"
join(34905.34955,35683.36040,36863.36942)
/gene="P0463F06.9"
/note="hypothetical protein"
/codon_start=1
/protein_id="BAB17123.1"
/db_xref="GI:11034598"
/translation="MSRVPYTSRGFVYSRVNKKARRSGRGKGNENAVHGPSKVDG
DDGKDRSGRRRRGSGRGRRRSGRRARRGGGRDGDAAKPEATPGRETVPASEK
RPEVVDGGERERRRLDSEGEKARQRVEMGEGERITPRQIRTEVYASAKASAGVC
ID"
39216.39842
/gene="P0463F06.10"
39216.39842
/gene="P0463F06.10"
/note="hypothetical protein
similar to Oryza sativa chromosome 6, P0675A05.3"
/codon_start=1
/protein_id="BAB17123.1"
/db_xref="GI:11034599"
/translation="MARGSALLDGSVLPSPRIYSEKQAGLPFRREMESATGREIVMLG
EARPAPDYPGRSTIFLPEFAMAGLVPPFSSFFMDVLEFYDLQMAHLTPNAVMTLAIFAH
LCMEFTGVNPSLRLEFMFPTVQSVSPSVVGGCYFLRGPVLNRYLPCVLRKKWDWK
SDMEFTPLADEAHLRLPSQPPVQASSWRAPVDLVSREVPPLPCSKFIK"
complement(join(40032.40444,40524.40710,40837.41124,
41213.41293))
/gene="P0463F06.11"
complement(join(40032.40444,40524.40710,40837.41124,
41213.41293))
/gene="P0463F06.11"
/note="hypothetical protein"

gene
CDS
alignment_scores:
Quality: 162.00 Length: 197
Ratio: 1.361 Gaps: 9
Percent Similarity: 60.406 Percent Identity: 23.858

alignment_block:
US-09-805-919-1 x AP002867/rev ..
Align seg 1/1 to reverse of: AP002867 from: 1 to: 144322

29 CysValProSerSerCysGlyAspIle..GlnIleLysPheProPheAr 44
||| ||| :||| ||| :||| :||| :||| :||| :||| :||| :|||
86765 TCCCTCCGTCACCTGCTGTAATCTTAGCAGCGTGTCTCTCTTCCG 86716

44 gLeuArgThrAspProGlnHisCysGlyArgArgGlyTyrGlnLeuAspC 61
| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
86715 CCGGCGAGGTGACCCATCTGAGTGGGAGTCTGTCTGACGAGCTCAATT 86666

61 ySgIAsnAsnGlnThValPheAsnTyrLysSerArgIlePheAspVal 77
|| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86665 GTGCCGATGACAAAGGCTACAAATTCAGATTGACAAACGGAACATATGTG 86616

78 GlnGluIleAsnTyrArgSerTyrSerIleArgLeuLeuAspProGlyLe 94
||| ||| ||| :||| :||| :||| :||| :||| :||| :||| :|||
86615 ACGGGCATCAACTACATGATTCTACCTTCTTGATCGTTGATGCTGACAT 86566

94 uAsnAspGlnArgGluAsnCysThrValPheProAsnHisArgAlaSerT 111
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86565 CTCGATTCACTAGCAAGTGCCTCTT.....CTTCGAT 86531

111 yAspAlaMetThrSerGlnIlePheGlu.....TrpValArgValAsn 125
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86530 GGAATTCCTTCGAGAGCCCAAGGTTCCAGGTTCTATGAGCTAGATA... 86484

126 AsnAspIleAsnTyrValAsnCysLeuAlaProIleAsnSerSer..... 140
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
```

```
86483 .....GCTAGCTTTGTCAAATTGTTCCGACGAAATAAGATAGTAGCTG 86440
141 ...GlnTyrIleProThrSerPheCysSerLysAsnSerThrGlyPhe 156
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86439 GATTATGATACATGCTTCGCT...TGCCCTGAGCACACGAGATCCTTCG 86393

156 eTyrLeuValIleArgGlu.....IleLeuGlnAlaSerAspLeu 169
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86392 TTTATGTTCTTAACGTGCCCAAAATCTGCTTTATTCAGAGCCTTGAGCT 86343

170 AlaGlyGlyCysArgValGlnThrValAlaTrp.....SerSerAl 183
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86342 TCTTGTGGGTACCTAGCCCATGACTCTTTGGGTGTCGGACAGTGTAGC 86293

183 aPro.....GlyIleSerSerAsnLysSerSerThrLeuSerSerT 197
||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86292 ACCAGTAATGAAGCTTAAGCTTAAGCTATGCAGACGTAGTAAATCTTA 86243

197 hrHisGlnGlyLeuAlaTyrGlyPheGlnLeuSerTrpLys 210
: ||| :||| :||| :||| :||| :||| :||| :||| :||| :|||
86242 TGAGCAAGCATTCAGAGTCCAAATTCTCCTTCAGATATTCCT 86202

seq_name: gb_pl:AF238473
seq_documentation_block:
LOCUS AF238473 6000 bp DNA linear PLN 21-JUN-2000
DEFINITION Oryza sativa receptor-like kinase (RLG12) gene, complete cds.
ACCESSION AF238473
VERSION AF238473.1 GI:8575479
KEYWORDS
SOURCE
ORGANISM
Oryza sativa.
Oryza sativa
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
Erihartoideae; Oryzeae; Oryza.
REFERENCE
1 (bases 1 to 6000)
Yun,C.-H., Lee,G.-R. and Kim,H.-I.
Molecular analysis of receptor-like kinase in rice
Unpublished
JOURNAL
2 (bases 1 to 6000)
Yun,C.-H., Lee,G.-R. and Kim,H.-I.
Direct Submission
REFERENCE
Submitted (24-FEB-2000) Cytoogenetics, National Institute of
Agricultural and Technology, Seodundong 249, Suwon 441-707, Korea
JOURNAL
FEATURES
source
1.6000
/organism="Oryza sativa"
/variety="Nipponbare"
/db_xref="taxon:4530"
/chromosome="1"
join(<1159..1931,4584..5670)
/gene="RLG12"
/product="receptor-like kinase"
<1159..5670
/gene="RLG12"
join(1159..1931,4584..5670)
/gene="RLG12"
/note="protein kinase"
/codon_start=1
/product="receptor-like kinase"
/protein_id="AAF78017.1"
/db_xref="GI:8575480"
/translation="MEAAALSTVLVLYALVYADADHHVVRVQRRHOCQPFSCGHLQN
VSPFRRRGDPDRSRGCFRSYELDSSDDGKATIRINTGKYVSSINYTASIFWYVDA
NLQDDANSSCPDRSDQLPYVGLGIPGSHDSWDLGLDMITWGVFVNGSOELITNSSS
KYLIVYHPVDCLTALSEFYVYMLPFSVLVSVEDIIEPSCGYLMTPIGGGWSVNLPL
EYTSYAEVVKSMRGFAVKFPTFWLWFGLIKKLKNSVSIHQAGHTTYCIAFISL
FVCSAMOVCVRTVGRDDESYSGOGYGSVYGVFLPGDVYVAIKMLDGNNGEDE
ISEVATIGRIHINNVRLVGFCSSEMRALVYEMPRGSLNKYIFSSERSFSWDLNE
IALGIARGINYLHQGCEMOILIFDIRDNILLDNFVSKYVDEGLAKLYPREKSFVSD
RALRGTVGIAPENYSRSFGVYISDKSDVYSFGMLLLEMAGRRNADPNANSNBSRAY
PAMVVDOLIADOQVDEISNVADMHLEKRLCLVGIWCIQMSHSDRPTMSEAIEMLEGD
VDALQVPLRPFCDGDGIGNGMPPQYMDSYFHSSELTATSEEDDGIAEIASS"
```



```

9728 AATTGTAGCTACAGTCCGAGGTTTCAACTCTCTTGCAACAACCTCAGCCA 9777
      :::::
65 nThrValPheAsnTyr...LysSerArgIlePheAspValGlnGluIleA 81
      |||:::
9778 AACCATTTGAAAGCTCCCAATTCAGGGGACTTAATTGTGAAAAAGCAATG 9827
      81 snTyArgSerTyrSerIleArgLeuLeuAspProGlyLeuAsnAspGln 97
      ::|||:::
9828 ATTACAGAGCCCAACCAATCCGAGTCAACGATCA..... 9862
      98 ArgIuAsnCysThrValPheProAsnHisArgAlaSerTyrAspAlaMe 114
      |||:::
9863 ...GAAGTTGT.....CTACCGAAC..... 9880
      114 tThrSerGlnIlePheGluTrpValArgValAsnAsnAspIleAsnTyrV 131
      |||:::
9881 .....CGGTTTCATGCACGATTGGAACCTCT 9905
      131 alaAsnCys.....LeuAlaProIleAsnSerSerGlnTyrIlePro 144
      ::|
9906 CCGGTTCAACCTTTCAGGCTAAACCCGATTATATACCAAGCTATTACAAAT 9955
      ::|
145 ThrSerPhe.....CysSerLysAsnSerThrGlyPheSerTyrLeuVa 159
      ::|||
9956 CTCACCTTCCTTCGTTGCCCAACAACCTCGACCGGT..... 9991
      159 lIleArgGluIleLeuGlnAlaSerAspLeuAlaGlyLysArgValG 176
      9991 ..... 9991
      176 luthrValAlaIleTrpSerSerAlaProGlyIleSer....SerAsn.Ly 190
      |||:::
9992 .....TCATCGTCTTTCGGTTTAGCTTGTGAGCAATATAA 10027
      190 sSerSerThrLeuSerSerThrHisGlnGlyLeuAlaTyrGlyPheGluL 207
      |:::
10028 GGGAAAAACA....AAACTCATCAGACACGACGACATTAGCTCCTCAAT 10071
      207 euSerTrpLysArg.....AsnLeuLeuCysArgAsnCysAspArgSer 221
      |||
10072 GGTCCTGGTATCGTGGCTCGGCTATTTGCTGTCACGTGCGCAACGAT 10121
      222 .ArgGlyGlyGluCysThrIleGluGluAsnSerAspArgAlaThrCysA 238
      |||
10122 GCGAGGTGATTGTTG.....GGTCTGTGTT 10147
      238 rGTYrTrpCysLysGluAspIleHisValSerLysLeu 250
      |||
10148 CCTATTTCGTGGCCGAGATGCCCATGTGCCCTTCTG 10185
```

